B89

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2002 (12.12.2002)

PCT

(10) International Publication Number
WO 02/099982 A2

(51) International Patent Classification[7]: H04B

(21) International Application Number: PCT/US02/06327

(22) International Filing Date: 1 March 2002 (01.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/272,803 1 March 2001 (01.03.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/272,803 (CIP)
Filed on 1 March 2001 (01.03.2001)

(71) Applicant *(for all designated States except US)*: **ILLUMINA, INC.** [US/US]; 9885 Towne Centre Drive, San Diego, CA 92121-1975 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DICKINSON, Todd** [US/US]; 3435 Lebon Drive #1133, San Diego, CA 92122 (US). **OLIPHANT, Arnold** [US/US]; 15518 Markar Road, Poway, CA 92064-2315 (US).

(74) Agents: **BREZNER, David, J.** et al.; Flehr, Hohbach, Test, Albritton & Herbert LLP, Suite 3400, Four Embarcadero Center, San Francisco, CA 94111-4187 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/099982 A2

(54) Title: METHODS FOR IMPROVING SIGNAL DETECTION FROM AN ARRAY

(57) Abstract: The invention relates methods of improving signal detection from an array and methods for background subtraction in an array. The invention provides for novel array compositions including arrays with wells with different shapes, or surfaces coated with reflective or selectively absorptive coatings. In addition, the array include a signal transducer element.

METHODS FOR IMPROVING SIGNAL DETECTION FROM AN ARRAY

This application claims the benefit of U.S.S.N. 60/272,803, which is hereby expressly incorporated by reference in its entirety.

FIELD OF THE INVENTION

The invention relates methods of improving signal detection from an array and methods for background subtraction in an array while maintaining optimal humidity levels. The invention provides for novel array compositions including arrays with wells with different shapes, surfaces coated with reflective or selectively absorptive coatings, buffering salts, bead spacing, and sub-pixel resolution of individual wells. In addition, the array include a signal transducer element.

BACKGROUND OF THE INVENTION

There are a number of assays and sensors for the detection of the presence and/or concentration of specific substances in fluids and gases. Many of these rely on specific ligand/antiligand reactions as the mechanism of detection. That is, pairs of substances (i.e. the binding pairs or ligand/antiligands) are known to bind to each other, while binding little or not at all to other substances. This has been the focus of a number of techniques that utilize these binding pairs for the detection of the complexes. These generally are done by labeling one component of the complex in some way, so as to make the entire complex detectable, using, for example, radioisotopes, fluorescent and other optically active molecules, enzymes, etc.

Of particular use in these sensors are detection mechanisms utilizing luminescence or fluorescence. Recently, the use of optical fibers and optical fiber strands in combination with light absorbing dyes for chemical analytical determinations has undergone rapid development, particularly within the last decade. The use of optical fibers for such purposes and techniques is described by Milanovich et al., "Novel Optical Fiber Techniques For Medical Application", Proceedings of the SPIE 28th Annual International Technical Symposium On Optics and Electro-Optics, Volume 494, 1980; Seitz, W.R.,

"Chemical Sensors Based On Immobilized Indicators and Fiber Optics" in *C.R.C. Critical Reviews In Analytical Chemistry*, Vol. 19, 1988, pp. 135-173; Wolfbeis, O.S., "Fiber Optical Fluorosensors In Analytical Chemistry" in *Molecular Luminescence Spectroscopy, Methods and Applications* (S. G. Schulman, editor), Wiley & Sons, New York (1988); Angel, S.M., *Spectroscopy 2* (4):38 (1987); Walt, et al., "Chemical Sensors and Microinstrumentation", *ACS Symposium Series*, Vol. 403, 1989, p. 252, and Wolfbeis, O.S., *Fiber Optic Chemical Sensors*, Ed. CRC Press, Boca Raton, FL, 1991, 2nd Volume.

When using an optical fiber in an *in vitro/in vivo* sensor, one or more light absorbing dyes are located near its distal end. Typically, light from an appropriate source is used to illuminate the dyes through the fiber's proximal end. The light propagates along the length of the optical fiber; and a portion of this propagated light exits the distal end and is absorbed by the dyes. The light absorbing dye may or may not be immobilized; may or may not be directly attached to the optical fiber itself; may or may not be suspended in a fluid sample containing one or more analytes of interest; and may or may not be retainable for subsequent use in a second optical determination.

Once the light has been absorbed by the dye, some light of varying wavelength and intensity returns, conveyed through either the same fiber or collection fiber(s) to a detection system where it is observed and measured. The interactions between the light conveyed by the optical fiber and the properties of the light absorbing dye provide an optical basis for both qualitative and quantitative determinations.

Of the many different classes of light absorbing dyes which conventionally are employed with bundles of fiber strands and optical fibers for different analytical purposes are those more common compositions that emit light after absorption termed "fluorophores" and those which absorb light and internally convert the absorbed light to heat, rather than emit it as light, termed "chromophores."

Fluorescence is a physical phenomenon based upon the ability of some molecules to absorb light (photons) at specified wavelengths and then emit light of a longer wavelength and at a lower energy. Substances able to fluoresce share a number of common characteristics: the ability to absorb light energy at one wavelength λ_{ab}; reach an excited energy state; and subsequently emit light at another light wavelength, λ_{em}. The absorption and fluorescence emission spectra are individual for each fluorophore and are often graphically represented as two separate curves that are slightly overlapping. The same fluorescence emission spectrum is generally observed irrespective of the wavelength of the exciting light and, accordingly, the wavelength and energy of the exciting light may be varied within limits; but the light emitted by the fluorophore will always provide the same emission spectrum. Finally, the strength of the fluorescence signal may be measured as the quantum yield of light emitted. The fluorescence quantum yield is the ratio of the number of photons emitted in comparison to the number of photons initially absorbed by the fluorophore. For more detailed information regarding each of these characteristics, the following references are recommended: Lakowicz, J. R., *Principles of*

Fluorescence Spectroscopy, Plenum Press, New York, 1983; Freifelder, D., Physical Biochemistry, second edition, W. H. Freeman and Company, New York, 1982; "Molecular Luminescence Spectroscopy Methods and Applications: Part I" (S.G. Schulman, editor) in *Chemical Analysis*, vol. 77, Wiley & Sons, Inc., 1985; *The Theory of Luminescence*, Stepanov and Gribkovskii, Iliffe Books, Ltd., London, 1968.

In comparison, substances which absorb light and do not fluoresce usually convert the light into heat or kinetic energy. The ability to internally convert the absorbed light identifies the dye as a "chromophore." Dyes which absorb light energy as chromophores do so at individual wavelengths of energy and are characterized by a distinctive molar absorption coefficient at that wavelength. Chemical analysis employing fiber optic strands and absorption spectroscopy using visible and ultraviolet light wavelengths in combination with the absorption coefficient allow for the determination of concentration for specific analyses of interest by spectral measurement. The most common use of absorbance measurement via optical fibers is to determine concentration which is calculated in accordance with Beers' law; accordingly, at a single absorbance wavelength, the greater the quantity of the composition which absorbs light energy at a given wavelength, the greater the optical density for the sample. In this way, the total quantity of light absorbed directly correlates with the quantity of the composition in the sample.

Many of the recent improvements employing optical fiber sensors in both qualitative and quantitative analytical determinations concern the desirability of depositing and/or immobilizing various light absorbing dyes at the distal end of the optical fiber. In this manner, a variety of different optical fiber chemical sensors and methods have been reported for specific analytical determinations and applications such as pH measurement, oxygen detection, and carbon dioxide analyses. These developments are exemplified by the following publications: Freeman, et al., *Anal Chem.* 53:98 (1983); Lippitsch et al., *Anal. Chem. Acta.* 205:1, (1988); Wolfbeis et al., *Anal. Chem.* 60:2028 (1988); Jordan, et al., *Anal. Chem.* 59:437 (1987); Lubbers et al., *Sens. Actuators* 1983; Munkholm et al., *Talanta* 35:109 (1988); Munkholm et al., *Anal. Chem.* 58:1427 (1986); Seitz, W. R., *Anal. Chem.* 56:16A-34A (1984); Peterson, et al., *Anal. Chem.* 52:864 (1980): Saari, et al., *Anal. Chem.* 54:821 (1982); Saari, et al., *Anal. Chem.* 55:667 (1983); Zhujun et al., *Anal. Chem. Acta.* 160:47 (1984); Schwab, et al., *Anal. Chem.* 56:2199 (1984); Wolfbeis, O.S., "Fiber Optic Chemical Sensors", *Ed. CRC Press, Boca Raton, FL*, 1991, 2nd Volume; and Pantano, P., Walt, D.R., *Anal. Chem.*, 481A-487A, Vol. 67, (1995).

More recently, fiber optic sensors have been constructed that permit the use of multiple dyes with a single, discrete fiber optic bundle. U.S. Pat. Nos. 5,244,636 and 5,250,264 to Walt, *et al.* disclose systems for affixing multiple, different dyes on the distal end of the bundle, the teachings of each of these patents being incorporated herein by this reference. The disclosed configurations enable separate optical fibers of the bundle to optically access individual dyes. This avoids the problem of deconvolving the separate signals in the returning light from each dye, which arises when the signals

from two or more dyes are combined, each dye being sensitive to a different analyte, and there is significant overlap in the dyes' emission spectra.

U.S.S.N.s 08/818,199 and 09/151,877 describe array compositions that utilize microspheres or beads on a surface of a substrate, for example on a terminal end of a fiber optic bundle, with each individual fiber comprising a bead containing an optical signature. Since the beads assemble randomly onto the surface, a unique optical signature is needed to "decode" the array; i.e. after the array is made, a correlation of the location of an individual site on the array with the bead or bioactive agent at that particular site can be made. This means that the beads may be randomly distributed on the array, a fast and inexpensive process as compared to either the *in situ* synthesis or spotting techniques of the prior art. Once the array is loaded with the beads, the array can be decoded, or can be used, with full or partial decoding occurring after testing, as is more fully outlined below.

Arrays, which perform individual assays in parallel, are developed to achieve the high-throughput required for large-scale genetic analysis. The spacing between test sites in an array defines the array's density. Higher density increases parallel processing. In addition to increasing the throughput, higher density reduces the required sample volume, and thereby lowers costs. Arrays offer parallel processing by performing multiple assays per sample simultaneously. However, current array formats lack the ability to test multiple samples simultaneously, one more level of parallel processing necessary for large-scale genetic analysis. In addition, most array technologies also have limited applications outside of SNP genotyping and gene expression profiling. Manufacturing limitations have further prevented arrays from reaching their full potential. Alternative methodologies for manufacturing arrays include mechanical deposition, inkjet printing and photolithography, each with its own set of limitations.

Accordingly, it is an object of the present invention to provide methods and compositions for improving the signals detected from arrays.

SUMMARY OF THE INVENTION

Accordingly, the invention provides a composition that includes a substrate with a surface comprising discrete sites, a reflective coating on the surface, and a population of microspheres distributed on the substrate. The microspheres comprise at least a first and a second subpopulation. Generally at least one subpopulation comprises a bioactive agent.

In addition, the invention provides a composition wherein the substrate comprises a first and a second surface, wherein the first surface comprises the discrete sites, and the reflective coating is on the second surface. The population of microspheres are distributed on the first surface.

In addition the invention provides a method of making a reflective array. The method includes providing a substrate with a surface comprising discrete sites, applying to the surface a coating of reflective material and distributing microspheres on the surface.

In addition, the invention provides a method, wherein the substrate comprises a first and a second surface, wherein the first surface comprises discrete sites, the reflective material is on the second surface and the microspheres are distributed on the first surface.

In addition the invention includes a method comprising providing a pre-formed unitary fiber optic bundle comprising a proximal and a distal end, the distal end comprising plurality of discrete sites comprising a population of microspheres, the population comprising at least first and second subpopulations, and imaging the fiber optic bundle from the distal end. A reflective coating may be applied to either the distal end or the proximal end of the fiber optic bundle.

In addition the invention provides an array composition comprising a substrate with a surface comprising discrete sites comprising alternatively shaped wells. The wells may contain a cross section that is shaped as a square, a hexagon, a star, a triangle, a pentagon or an octagon.

Accordingly the invention provides a method comprising providing a substrate with a plurality of discrete sites, the sites comprising alternatively shaped wells and a population of microspheres, the population comprising at least first and second subpopulations, and imaging the substrate.

In addition the invention provides an array composition comprising a substrate with a surface comprising discrete sites and a population of microspheres distributed on the substrate, wherein the microspheres comprise a bioactive agent and a signal transducer element.

Accordingly the invention provides a method of detecting a non-labeled target analyte in a sample comprising providing a substrate with a plurality of discrete sites, distributing on the sites a population of microspheres comprising a bioactive agent and a signal transducer element, contacting the substrate with the sample, whereby upon binding of the target analyte to the bioactive agent, a signal from the signal transducer element is altered as an indication of the presence of the target analyte.

In addition, the new profile shape of the wells improves bead retention and bead array uniformity in various substrates.

In addition the invention provides a method of detecting a chiral molecule in a sample comprising providing a substrate with a surface comprising at least first and second discrete sites at least first and second bioactive agents attached to the first and second discrete sites respectively, contacting the substrate with the sample, illuminating the substrate with polarized light, and detecting rotation of the

light in at least one of the first and second discrete sites as an indication of the presence of the chiral molecule.

In addition, the invention provides a method of using the ratio of beads with bioactive agents to blank beads to achieve proper density of beads.

In addition the invention provides a method of determining the location of a microsphere in an array comprising providing a substrate with a first surface comprising at least a first and a second discrete site, wherein the first discrete site comprises a microsphere, but the second discrete site does not comprising a microsphere, illuminating the substrate and detecting illumination of the substrate, whereby reduced illumination at the first discrete site relative to the second discrete site provides an indication of the presence of the first microsphere in the first discrete site.

In addition the invention provides a method of increasing signal output from an array comprising providing a substrate with a surface comprising at least first and second discrete sites and at least first and second labels attached to the first and second discrete sites respectively, cooling the substrate to at least below room temperature and detecting a signal from the first and second labels, whereby the signal is increased relative to a signal obtained from a substrate that is not cooled.

In addition the invention provides a method for background signal subtraction in an array comprising providing a substrate with a surface comprising at least first and second discrete sites and at least first and second labels attached to the first and second discrete sites respectively, detecting the signal from the first and second discrete sites in a plurality of different emissions, and subtracting the lowest signal from each of the first and second discrete sites from the remaining signals from the first and second discrete sites, respectively.

In addition the invention provides a method of correcting image non-uniformity comprising providing a substrate with a surface comprising at least first and second discrete sites, at least first and second labels attached to the first and second discrete sites respectively and at least a first internal reference point of known signal intensity, detecting a first and second signal from the first and second labels, respectively, detecting a signal from the internal reference point, and determining the variation between the signal from the internal reference point and the known signal intensity of the internal reference point as an indication of said image non-uniformity.

In addition the invention provides a method of detecting a target analyte in a sample comprising providing an array comprising a substrate with a surface comprising discrete sites, a reflective coating on said surface, and a population of microspheres distributed on the substrate. The microspheres compriset least a first and a second subpopulation each comprising a different bioactive agent. The method further includes contacting the array with the sample, such that the target analyte binds to at

least one of the bioactive agents and detecting the presence of the target analyte. In a preferred embodiment the target analyte is labeled.

In addition the invention provides a method of detecting a target analyte in a sample comprising providing an array comprising a substrate with a surface comprising discrete sites comprising alternatively shaped wells and a population of microspheres distributed on the substrate. The microspheres comprise at least a first and a second subpopulation each comprising a different bioactive agent. The method further includes contacting the array with the sample, such that the target analyte binds to at least one of the bioactive agents and detecting the presence of the target analyte.

In addition the invention provides a method of detecting a target analyte in a sample comprising providing a substrate with a surface comprising at least first and second discrete sites and a population of microspheres distributed on the substrate, wherein the microspheres comprise at least a first and a second subpopulation each comprising a different bioactive agent, contacting the substrate with the sample, such that the target analyte binds to at least one of the bioactive agents. The invention further includes cooling the substrate to at least below room temperature and detecting a signal, whereby the signal is increased relative to a signal obtained from a substrate that is not cooled.

In addition, the invention provides for the use of hygroscopic salt solutions to allow for careful control of the volumes during DNA hybridization to the beads. The advantage of being able to use very small volumes without losing too much water by evaporation or gaining too much water by adsorption is desirable.

BRIEF DESCRIPTIONS OF THE FIGURES

Figure 1 depicts the average signals and backgrounds fo a subset of beads imaged from the proximal and distal ends of a randomly assembled fiber optic array.

Figure 2 depicts signals and backgrounds from fluorescent beads in palladium-coated vs. uncoated microwell arrays. 2a) comparison of signals and backgrounds from a single bead type (fluorescein-labelled silica) assembled in a Pd-coated vs. uncoated etched fiber array. Fluorescence image of beads in b) uncoated etched fiber (signal to background of 2.47), and c) Pd-coated etched fiber (signal to background of 28.78).

Figure 3 depicts microspheres deposited in wells with concave shaped bottoms.

Figure 4 depicts signal output from beads distributed in wells of varying shapes. a) depicts standard well shape obtained from etching a fiber optic core; b) depicts a well shaped with sloped well walls; c) depicts rounded well

Figure 5 depicts the signal intensities of the same batch of dyed beads assembled into wells of different shapes.

Figure 6 depicts variation of hybridization signal as a function of well shape.

The data demonstrate that well shape has a significant impact on the quantity of fluorescence that is collected from a bead in that well.

Figure 7 depicts a micrograph of a silica bead distributed in a star-shaped well. The bead is essentially held at four contact points within a star-shaped well, while the rest of the circumference of the bead is unhindered, free from contact with the well substrate.

Figure 8 depicts intensity variation between dyed beads in wells vs. on adjacent flat surfaces.

Figure 9 depicts an analytical method for detecting an unlabeled target analyte. Upon binding of the unlabeled analyte to the bioactive agent coupled to a signal transducer, a property of the signal transducer, such as fluorescence, changes resulting in a changed signal.

Figure 10 depicts an analytical method for detecting an unlabeled target analyte. Upon binding of the unlabeled analyte to the bioactive agent coupled to a bead by a signal transducer molecule, a property of the signal transducer, such as fluorescence, changes resulting in a changed signal.

Figure 11 depicts an analytical method for detecting an unlabeled target analyte. An unlabeled analyte binds to a recognition sequence of a bioactive agent. The bioactive agent is bound to the bead by associating with a generic tag on the signal transducer that is attached to the bead. Upon binding of the target analyte to the bioactive agent, a property of the signal transducer, such as fluorescence, changes resulting in a changed signal.

Figure 12 depicts signal intensity of beads that have been exposed to cold temperatures.

Figure 13 depicts a method of background subtraction for identifying the location of a bead in an array without the use of labeled beads. Figure 13A: Shows a zoomed in image of an Bead Array™ taken as described in this invention where one of the cores has a bead present (A) and appears relatively dimmer than other empty cores in the array. Figure 13B: Shows the same Bead Array™ after

hybridization to a fluorescent labeled oligonucleotide complement demonstrating that the core has a bead present.

Figure 14 depicts a schematic of a reflective film or coating scenario for bottom-coating a substrate.

Figure 15 depicts a bead-in-a-well within a depression that contains bound enzyme at sites surrounding the bead. The signaling molecule diffusing from the bead can encounter an enzyme within a defined region (i.e. the depression) and thereby generate light. The shape of the depression as shown is arbitrary, and can be adjusted to optimize efficiency of light generation. In addition the shape of the well in which the bead resides also can be adjusted to optimize signal generation as described in more detail herein.

Figure 16a) depicts a fiber-optic etched well profile.

Figure 16b) depicts a microfabricated plastic well profile.

Figure 17 depicts the schematic of the auto-centering effect of angled wells.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to the use of randomly ordered arrays comprising a bead-based analytic chemistry system in which beads, also termed microspheres, carrying different chemical functionalities are distributed on a substrate comprising a patterned surface of discrete sites that can bind the individual microspheres. The beads are generally put onto the substrate randomly, i.e. each bead goes down arbitrarily or indiscriminately on to a site. This allows the synthesis of the candidate agents (i.e. compounds such as nucleic acids and antibodies) to be divorced from their placement on an array, i.e. the candidate agents may be synthesized on the beads, or on a different substrate and then put onto the beads, and then the beads are randomly distributed on a patterned surface.

However, the random placement of the beads means that all or part of the array must be "decoded" after synthesis; that is, after the array is made, a correlation of the location of an individual site on the array with the bead or candidate agent at that particular site can be made. This encoding/decoding can be done in a number of ways, as is generally described in 60/090,473; 09/189,543; 08/944,850; 08/818,199; 09/151,877; and 08/851,203, all of which are hereby expressly incorporated by reference in their entirety. These methods include: (1) "encoding" the beads with unique optical signatures, generally fluorescent dyes, that can be used to identify the chemical functionality on any particular bead; (2) using a decoding binding ligand (DBL), generally directly labeled, that binds to either the bioactive agent or to identifier binding ligands (IBLs) attached to the beads; (3) the use of positional decoding, for example by either targeting the placement of beads (for example by using

photoactivatible or photocleavable moieties to allow the selective addition of beads to particular locations), or by using either sub-bundles or selective loading of the sites, as are more fully outlined below; (4) the use of selective decoding, wherein only those beads that bind to a target are decoded; or (5) combinations of any of these. In some cases, as is more fully outlined below, this decoding may occur for all the beads, or only for those that bind a particular target analyte. Similarly, this may occur either prior to or after addition of a target analyte.

This means that the beads may be randomly distributed on the array, a fast and inexpensive process as compared to either the *in situ* synthesis or spotting techniques of the prior art.

Once the identity (i.e. the actual agent) and location of each microsphere in the array has been fixed, the array is exposed to samples containing the target analytes, although as outlined below, this can be done prior to, during or after the assay as well. The target analytes will bind to the bioactive agents as is more fully outlined below, and results in a change in an optical signal of a particular bead.

The present invention is directed to compositions and methods that improve signal detection or assay performance in an array. That is, in the broadest sense, the invention provides an array composition comprising a substrate with a surface comprising discrete sites, wherein the discrete sites are geometrically shaped wells. That is, wells within a microarray are formed in specific shapes depending on the sample to be examined, assay to be performed or size or shape of microsphere to be distributed in the well.

In a preferred embodiment, the microwells are designed for use with beads, for example in Illumina's BeadArrays™, which possess profiles that are ideally suited for retaining beads in an ordered or randomly-ordered array. Following an acid etch step, fiber-optic substrates frequently generate microwells with convex bottoms (Figure 16a). Wells with other shapes also find use in arrays. When using an optical fiber in an *in vitro/in vivo* sensor, optimizing well geometry for bead retention and array spacing uniformity facilitates detection. A new profile shape of cell wells improve bead retention and bead array uniformity.

While it is possible to alter glass compositions and the fiber drawing process such that the etched wells are concave at the well-bottom surface, a preferred alternative is to microfabricate the wells in other materials (e.g. plastic and other polymers) using photolithography and/or other processes that allow one to produce virtually any desired well shape and profile.

In a preferred embodiment, microspheres are anchored within microwells through van der Waals interactions between the bead and the walls of the well. In general, bead arrays exhibit the best retention when the beads are matched most closely in size to the diameter of the well into which they are assembled. This suggests that retention is directly related to the degree of contact a bead has with

the walls of its well. Thus any design that maximizes surface area contact between the bead and the well may lead to further improvements in bead retention.

In a preferred embodiment, the present invention, therefore, is to microfabricate wells with concave "floors" that come as close to matching the shape of the bead as possible (Figure 16b). An additional design feature is to add a slight draft angle or taper to the well walls; this serves to: 1) improve filling efficiency by providing slightly wider "mouths" to the wells and thus improve the probability and speed with which wells are filled by beads, and 2) improve the spacing uniformity of the beads in the array since the angled wall will automatically force the beads to fall down into the well until they "plug" in the center of the well. BeadArrays™ platforms, planar or fiber based arrays, may utilize this well profile.

In an alternative embodiment, the invention is directed to a composition comprising a substrate that has reduced or diminished auto-fluorescence characteristics. In addition, the invention includes the use of non-fluorescent coatings on the substrate to reduce substrate fluorescence. In addition, the invention includes a composition comprising a reflective substrate. The reflective properties may be property of the substrate itself, or may be the result of a reflective coating on the substrate. In one embodiment the reflective property is the result of the shape of wells or discrete sites on the substrate.

In an alternative embodiment the invention is directed to a composition comprising a substrate comprising at least one assay location containing at least a first discrete site surrounded by functionalized sites for the attachment of detection moieties. In a preferred embodiment the substrate comprises a plurality of assay locations containing at least one discrete site or well surrounded by functionalized sites for the attachment of detection moieties or surounded by attachment moieties attached to the substrate.

Accordingly, the present invention provides random array compositions comprising at least a first substrate with a surface comprising individual sites. By "random" array herein is meant an array that is manufactured under conditions that results in the identification of the agent in at least some, if not all, of the sites of the array being initially unknown; that is, each agent is put down arbitrarily on a site of the array in a generally non-reproducible manner. What is important in random arrays, and what makes the present invention so useful, is that random arrays generally require at least one, and generally several "decoding" steps that produce data images that must be compared. In addition, while the techniques of the invention can be used on a variety of random arrays, the discussion below is directed to the use of arrays comprising microspheres that are laid down randomly on a surface comprising discrete sites. However, as will be appreciated by those in the art, other types of random arrays, i.e. those not containing beads, may also utilize the methods of the invention.

By "array" herein is meant a plurality of candidate agents in an array format; the size of the array will depend on the composition and end use of the array. Arrays containing from about 2 different bioactive

agents (i.e. different beads) to many millions can be made, with very large fiber optic arrays being possible. Generally, the array will comprise from two to as many as a billion or more, depending on the size of the beads and the substrate, as well as the end use of the array, thus very high density, high density, moderate density, low density and very low density arrays may be made. Preferred ranges for very high density arrays (all numbers are per cm^2) are from about 10,000,000 to about 2,000,000,000, with from about 100,000,000 to about 1,000,000,000 being preferred. High density arrays range about 100,000 to about 10,000,000, with from about 1,000,000 to about 5,000,000 being particularly preferred. Moderate density arrays range from about 10,000 to about 100,000 being particularly preferred, and from about 20,000 to about 50,000 being especially preferred. Low density arrays are generally less than 10,000, with from about 1,000 to about 5,000 being preferred. Very low density arrays are less than 1,000, with from about 10 to about 1000 being preferred, and from about 100 to about 500 being particularly preferred. In some embodiments, the compositions of the invention may not be in array format; that is, for some embodiments, compositions comprising a single bioactive agent may be made as well. In addition, in some arrays, multiple substrates may be used, either of different or identical compositions. Thus for example, large arrays may comprise a plurality of smaller substrates.

In addition, one advantage of the present compositions is that particularly through the use of fiber optic technology, extremely high density arrays can be made. Thus for example, because beads of 200 μm or less (with beads of 200 nm possible) can be used, and very small fibers are known, it is possible to have as many as 40,000 or more (in some instances, 1 million) different fibers and beads in a 1 mm^2 fiber optic bundle, with densities of greater than 25,000,000 individual beads and fibers (again, in some instances as many as 100 million) per 0.5 cm^2 obtainable.

By "substrate" or "solid support" or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of beads and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates is very large. Possible substrates include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon ®, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, optical fiber bundles, and a variety of other polymers. In general, the substrates allow optical detection and do not themselves appreciably fluoresce.

Generally the substrate is flat (planar), although as will be appreciated by those in the art, other configurations of substrates may be used as well; for example, three dimensional configurations can be used, for example by embedding the beads in a porous block of plastic that allows sample access to the beads and using a confocal microscope for detection. Similarly, the beads may be placed on

the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Preferred substrates include optical fiber bundles as discussed below, and flat planar substrates such as glass, polystyrene and other plastics and acrylics.

In a preferred embodiment the substrate is made of a substance with low intrinsic fluorescence. That is, one of the primary sources of background in microarray systems is the intrinsic fluorescence of the array substrate. Accordingly, a substrate material as described above that is opaque i.e. black, has reduced intrinsic fluorescence and is a preferred substrate. Without being bound by theory, it is thought that the reduced fluorescence is a result of efficient light absorption by the material.

In one embodiment, the substrate may be coated with a material that has reduced or diminished intrinsic fluorescence properties. That is, in this embodiment the fluorescence of the substrate is masked or covered by the application of a non-flourescent covering. The ability to apply a non-fluorescent coating over a patterned substrate may obviate the need to use materials with intrinsically low fluorescence, thereby broadening the scope of the materials available for generating an array. As depicted in Figure 2, a thin palladium film coated on an array surface increased or improved signal-to-background ratios.

An additional benefit to coating a microarray substrate material is that it becomes more efficient at signal collection as a result of signal reflection. That is, the optical signal of the bead itself is reflected thereby increasing the signal of the bead(s). There are a variety of coatings that find use in this invention. These include but are not limited to gold, silver, chromium, platinum or indium tin oxide.

In one embodiment, the substrate contains two surfaces. That is, for example, a fiber optic bundle contains a proximal and a distal end or a planar substrate contains a top and a bottom surface. Accordingly, in one embodiment, a reflective coating is applied to the surface that contains the discrete sites or wells. Alternatively, the reflective coating is applied to the surface of the substrate that does not contain the discrete sites or wells.

For example, when microspheres are distributed in the distal end of a fiber optic bundle, the reflective coating is applied to the proximal end. That is, the proximal surface of the fiber-optic based array instead of the distal end contains the reflective coating.

In an alternative embodiment, the substrate is a planar substrate, such as a slide or chip. In this embodiment, discrete sites or wells are in one surface of the substrate while the reflective coating is applied to the other surface. In a preferred embodiment, the substrate is transparent.

In one embodiment, the substrate is coated with a substance that is transparent, but becomes reflective upon activation with heat, light, chemicals, or other means. In a preferred embodiment a

silver ion solution is used to coat the surface, with subsequent treatment with a reducing agent and light to create a silver 'mirror' or reflective layer on the surface. While the creation of 'mirrored' surfaces by the reduction of silver ion is a well known phenomenon, the process has not previously been applied to arrays including fiber optic arrays.

In a preferred embodiment, the distal end of an etched fiber bundle is coated with a developable reflective surface treatment. such as, but not limited to the silver solution described above. Advantages of using such a coating on a fiber optic array includes: (1) reading the array from the proximal end, (2) developing the coating on the distal end, and (3) reading the array from the distal end. The reflective surface permits reading of the distal without optical noise arising from the fiber or impurities in the fiber. One application of this method would be to use the remote sensing property of proximal end reading for array decoding, then the low background property of distal end reading for sensitive analyses such as the detection of target analytes.

In a preferred embodiment, the invention provides a method of making an array with a developable reflective coating. The method includes providing an array substrate. The substrate can be any substrate as described herein, although fiber optic bundles are preferred. When fibers are used, the distal ends are preferably polished and etched together. The distal end of the etched fibers are then incubated in a solution containing, for example, but not limited to, silver ion, and then washed to remove excess silver ion. The distal end is placed into a suspension of beads and the array assembled.

Once formed, the array finds use in any assay as described herein. In a preferred embodiment, decoding of beads is accomplished via the proximal end of the fiber. The proximal end of the fibers are then blackened. The distal ends of the fibers are placed into a solution of a reducing agent (e.g., sodium thiosulfate) and illuminated; the mirror forms on the fiber under and around the beads. The array can then be analyzed from the distal end.

In one embodiment, by coating the surface of the substrate that is opposite to the surface to which microspheres are to be distributed, the retention of beads in the wells may be improved. That is, an advantage of coating the "bottom" of the substrate is that microwell array surface is not compromised by the coating procedure. By "bottom" surface is meant that surface of a substrate that is opposite to the side upon which microspheres are distributed.

As demonstrated in Figure 14, a reflective coating applied to the bottom of a substrate serves to recycle the light from both the microspheres and the excitation source.

In one embodiment the coating applied to the surface is a dielectric coating. In an alternative embodiment, the coating selectively absorbs certain wavelengths.

In an alternative embodiment, the surface of a substrate is rendered reflective by generating or treating it in such a way as to create a very smooth finish. That is, if the well interior surface becomes rough such that it is not reflective, it may be necessary to treat the array in a manner, such as gentle and/or partial melting by processes as are known in the art, to return the glossy or reflective finish to the substrate.

In an alternative embodiment, reflective coatings on the substrate are not required. In this embodiment, the substrate contains wells that are concave in shape. Wells are shaped as is known in the art by such methods as etching, imprinting, stamping, ablating and the like. As such, the concave wells act as mirrors at the bottom of each well (Figure 3) to reduce undesirable background light. Such a well design not only re-directs stray excitation light back to the bead to generate additional fluorescence, but it also would reflect fluorescence emission form the bead back into the collection optics for improved signal collection.

In addition, as is more fully outlined below, the substrate may include a coating, edging or sheath of material, generally detectable, that defines a substrate edge that may serve as one or more fiducials.

In a preferred embodiment, the substrate is an optical fiber bundle or array, as is generally described in U.S.S.N.s 08/944,850 and 08/519,062, PCT US98/05025, and PCT US98/09163, all of which are expressly incorporated herein by reference. Preferred embodiments utilize preformed unitary fiber optic arrays. By "preformed unitary fiber optic array" herein is meant an array of discrete individual fiber optic strands that are co-axially disposed and joined along their lengths. The fiber strands are generally individually clad. However, one thing that distinguished a preformed unitary array from other fiber optic formats is that the fibers are not individually physically manipulatable without intentionally treating the preformed unitary array with agents that separate them, for example treating a preformed array susceptible to acid with an acid such that the interstitial material is etched and thus the individual cores can be separated. However, absent these intentional treatments, one strand generally cannot be physically separated at any point along its length from another fiber strand.

At least one surface of the substrate is modified to contain discrete, individual sites for later association of microspheres. These sites may also be referred to in some embodiments as "features". These sites may comprise physically altered sites, i.e. physical configurations such as wells or small depressions in the substrate that can retain the beads, such that a microsphere can rest in the well, or the use of other forces (magnetic or compressive), or chemically altered or active sites, such as chemically functionalized sites, electrostatically altered sites, hydrophobically/ hydrophilically functionalized sites, spots of adhesive, etc.

The sites may be a pattern, i.e. a regular design or configuration, or randomly distributed. A preferred embodiment utilizes a regular pattern of sites such that the sites may be addressed in the X-Y

coordinate plane. "Pattern" in this sense includes a repeating unit cell, preferably one that allows a high density of beads on the substrate. However, it should be noted that these sites may not be discrete sites. That is, it is possible to use a uniform surface of adhesive or chemical functionalities, for example, that allows the attachment of beads at any position. That is, the surface of the substrate is modified to allow attachment of the microspheres at individual sites, whether or not those sites are contiguous or non-contiguous with other sites. Thus, the surface of the substrate may be modified such that discrete sites are formed that can only have a single associated bead, or alternatively, the surface of the substrate is modified and beads may go down anywhere, but they end up at discrete sites.

In a preferred embodiment, the surface of the substrate is modified to contain wells, i.e. depressions in the surface of the substrate. This may be done as is generally known in the art using a variety of techniques, including, but not limited to, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those in the art, the technique used will depend on the composition and shape of the substrate.

In a preferred embodiment, the wells of the substrate are shaped in discrete shapes. That is, the shape of the wells is distinct from the shape obtained by etching acid-soluble cores of an optical fiber as described in U.S.S.N. 09/151,877, and U.S.P.N. 6,023,4540, both of which are hereby expressly incorporated by reference. That is, the wells are non-cylinder shapes. Wells in a microarray format can be custom designed in size, shape, depth and profile to improve assay performance and/or signal output. Improvements can be realized throughout the various stages of an assay as described below, including enhanced fluorescence signal collection, better filling efficiency and bead retention, reduced inter-element cross-talk and improved hybridization kinetics.

Wells can be manufactured in an array through well-established microfabrication techniques as described herein and as are known in the art.

As described in detail in U.S.S.N. 09/151,877, and U.S.P.N. 6,023,4540, both of which are hereby expressly incorporated by reference, wells in a fiber optic bundle are formed by etching acid-soluble cores of a circular optical fiber. However, it is possible that the vertical wall of the well that results from the etching may not be optimal for direct imaging (Figure 4a). By direct imaging is meant directly imaging a microsphere in an array as opposed to imaging the microsphere through the opposite end of an optical fiber. Accordingly, in one embodiment, the invention provides a substrate with alternative well shapes. The well shapes may include a sloped or more gradual wall angle (Figure 4B), or a rounded wall or rounded interior configuration, i.e. a partial spherical configuration (Figure 4C) may be used. In some embodiments, it is thought that the alternative well shapes results in a larger area of the bead that is accessible for optical interrogation resulting in increased signal sensitivity.

In an additional embodiment the wells are shaped in different geometric shapes. By different geometric shapes is meant a shape other than circular. These include but are not limited to a square, hexagon, star, triangle, pentagon or octagon. Signal intensity of beads distributed in differently shaped wells is depicted in Figure 5. Different hybridization signals of beads in different shaped wells is depicted in Figure 6. As shown in Figures 5 and 6, well shape has a significant on the quantity of fluorescence that is collected from a bead in a well. That is, for example, star-shaped wells allowed for increased signal detection compared to other shapes examined. Without being bound by theory, it is thought that this result is a consequence of the manner in which a bead rests within a star-shaped well.

As shown in Figure 7, the bead contacts the star-shaped well circumferentially at essentially four points. By reducing the amount of contact area between the bead and the substrate, increased signal output from the bead is observed. That is, when a bead rests tightly in a spherical well, it is thought that emission fluors in the bottom half of the bead are not easily detected. One reason for this may be an increased absorbance of the fluors by the substrate. This results in reduced signal collection by the optical detection system. Accordingly, by reducing the contact area between the bead and the substrate, increased signal output or detection is observed from the array.

In addition, by distributing beads in geometrically shaped wells i.e. star-shaped, improved hybridization efficiency is observed as shown in Figure 6. That is, for example, with star-shaped wells, the areas of the well that are not in contact with the bead, serve as channels or inlets for the sample solution to interact with the beads or substances i.e. bioactive agents or oligonucleotide probes, on the beads.

Accordingly, in one embodiment the invention provides a method of increasing signal output from an array. The method includes providing a substrate with a plurality of alternatively shaped wells, distributing labeled microspheres in the wells and imaging the array.

In an additional embodiment the invention provides for method of decreasing the contact area between a bead and a substrate. The rest of the bead, therefore, is free from contact with the substrate allowing for increased bead surface area that can contact assay solutions.

In a preferred embodiment, physical alterations are made in a surface of the substrate to produce the sites. In a preferred embodiment, the substrate is a fiber optic bundle and the surface of the substrate is a terminal end of the fiber bundle, as is generally described in 08/818,199 and 09/151,877, both of which are hereby expressly incorporated by reference. In this embodiment, wells are made in a terminal or distal end of a fiber optic bundle comprising individual fibers. In this embodiment, the cores of the individual fibers are etched, with respect to the cladding, such that small wells or depressions are formed at one end of the fibers. The required depth of the wells will depend on the size of the beads to be added to the wells.

Generally in this embodiment, the microspheres are non-covalently associated in the wells, although the wells may additionally be chemically functionalized as is generally described below, cross-linking agents may be used, or a physical barrier may be used, i.e. a film or membrane over the beads.

In a preferred embodiment, the surface of the substrate is modified to contain chemically modified sites, that can be used to attach, either covalently or non-covalently, the microspheres of the invention to the discrete sites or locations on the substrate. "Chemically modified sites" in this context includes, but is not limited to, the addition of a pattern of chemical functional groups including amino groups, carboxy groups, oxo groups and thiol groups, that can be used to covalently attach microspheres, which generally also contain corresponding reactive functional groups; the addition of a pattern of adhesive that can be used to bind the microspheres (either by prior chemical functionalization for the addition of the adhesive or direct addition of the adhesive); the addition of a pattern of charged groups (similar to the chemical functionalities) for the electrostatic attachment of the microspheres, i.e. when the microspheres comprise charged groups opposite to the sites; the addition of a pattern of chemical functional groups that renders the sites differentially hydrophobic or hydrophilic, such that the addition of similarly hydrophobic or hydrophilic microspheres under suitable experimental conditions will result in association of the microspheres to the sites on the basis of hydroaffinity. For example, the use of hydrophobic sites with hydrophobic beads, in an aqueous system, drives the association of the beads preferentially onto the sites. As outlined above, "pattern" in this sense includes the use of a uniform treatment of the surface to allow attachment of the beads at discrete sites, as well as treatment of the surface resulting in discrete sites. As will be appreciated by those in the art, this may be accomplished in a variety of ways.

In one embodiment the substrate comprises a patterned substrate that includes sites or wells into which beads are distributed and also contains a pattern that surrounds the sites (Fig. I5). In a preferred embodiment, the region surrounding each site includes functionalized sites for the attachment of detection moieties. By "detection moieties" is meant any molecule that detects a substance and produces a detectable signal, either directly or indirectly. Examples of detection moieties include Fluorescence resonance energy transfer (FRET) molecules that may be directly or indirectly attached to the functionalized site, reporter molecules including dyes and enzymes. In a preferred embodiment the detection moiety includes reporter enzymes such as but not limited to firefly luciferase, and b-galactosidase, The functionalized sites can include any of the functionialities as described herein.

In a preferred embodiment the region around each discrete site or bead is used to immobilize one or more detection enzymes. Luciferase and/or sulfurylase are particularly preferred. They find particular use in pyrosequencing.

The region, or reaction well, is designed so that a signaling molecule produced by the primary reaction, that occurs at the discrete site or well, has a high probability of encountering bound enzyme and therefore generates a signal (see Figure 15).

In a preferred embodiment the invention provides an array of assay regions each containing a discrete site surrounded by sites for attachment of detection moieties. By assay region is meant a discrete locus surrounding a well that contains functionalized sites for attachment of detection moieties as described herein. In a preferred embodiment the assay regions are physically separated from one another. The separation can be any physical barrier that prevents solutions from contaminating adjacent regions.

In one embodiment each assay region contains a plurality of wells into which beads are distributed. Again the region surrounding at least some if not all of the wells contains functionalized sites for the attachment of detection moieties. As is appreciated by one of ordinary skill in the art, where the array is described as having functionalized sites for the attachment of a detection moiety, the array also can include detection moieties attached to the substrate.

In a preferred embodiment, the compositions of the invention further comprise a population of microspheres. By "population" herein is meant a plurality of beads as outlined above for arrays. Within the population are separate subpopulations, which can be a single microsphere or multiple identical microspheres. That is, in some embodiments, as is more fully outlined below, the array may contain only a single bead for each bioactive agent; preferred embodiments utilize a plurality of beads of each type.

By "microspheres" or "beads" or "particles" or grammatical equivalents herein is meant small discrete particles. The composition of the beads will vary, depending on the class of bioactive agent and the method of synthesis. Suitable bead compositions include those used in peptide, nucleic acid and organic moiety synthesis, including, but not limited to, plastics, ceramics, glass, polystyrene, methylstyrene, acrylic polymers, paramagnetic materials, thoria sol, carbon graphited, titanium dioxide, latex or cross-linked dextrans such as Sepharose, cellulose, nylon, cross-linked micelles and teflon may all be used. *"Microsphere Detection Guide"* from Bangs Laboratories, Fishers IN is a helpful guide.

The beads need not be spherical; irregular particles may be used. In addition, the beads may be porous, thus increasing the surface area of the bead available for either bioactive agent attachment or tag attachment. The bead sizes range from nanometers, i.e. 100 nm, to millimeters, i.e. 1 mm, with beads from about 0.2 micron to about 200 microns being preferred, and from about 0.5 to about 5 micron being particularly preferred, although in some embodiments larger or smaller beads may be used.

It should be noted that a key component of the invention is the use of a substrate/bead pairing that allows the association or attachment of the beads at discrete sites on the surface of the substrate, such that the beads do not move during the course of the assay.

Each microsphere comprises a bioactive agent, although as will be appreciated by those in the art, there may be some microspheres which do not contain a bioactive agent, depending the on the synthetic methods. By "candidate bioactive agent" or "bioactive agent" or "chemical functionality" or "binding ligand" herein is meant as used herein describes any molecule, *e.g.*, protein, oligopeptide, small organic molecule, coordination complex, polysaccharide, polynucleotide, etc. which can be attached to the microspheres of the invention. It should be understood that the compositions of the invention have two primary uses. In a preferred embodiment, as is more fully outlined below, the compositions are used to detect the presence of a particular target analyte; for example, the presence or absence of a particular nucleotide sequence or a particular protein, such as an enzyme, an antibody or an antigen. In an alternate preferred embodiment, the compositions are used to screen bioactive agents, i.e. drug candidates, for binding to a particular target analyte.

Bioactive agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Bioactive agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The bioactive agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Bioactive agents are also found among biomolecules including peptides, nucleic acids, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are nucleic acids and proteins.

Bioactive agents can be obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification and/or amidification to produce structural analogs.

In a preferred embodiment, the bioactive agents are proteins. By "protein" herein is meant at least two covalently attached amino acids, which includes proteins, polypeptides, oligopeptides and peptides.

The protein may be made up of naturally occurring amino acids and peptide bonds, or synthetic peptidomimetic structures. Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. For example, homo-phenylalanine, citrulline and norleucine are considered amino acids for the purposes of the invention. The side chains may be in either the (R) or the (S) configuration. In the preferred embodiment, the amino acids are in the (S) or L-configuration. If non-naturally occurring side chains are used, non-amino acid substituents may be used, for example to prevent or retard *in vivo* degradations.

In one preferred embodiment, the bioactive agents are naturally occurring proteins or fragments of naturally occurring proteins. Thus, for example, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of procaryotic and eukaryotic proteins may be made for screening in the systems described herein. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred.

In a preferred embodiment, the bioactive agents are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized bioactive proteinaceous agents.

In a preferred embodiment, a library of bioactive agents are used. The library should provide a sufficiently structurally diverse population of bioactive agents to effect a probabilistically sufficient range of binding to target analytes. Accordingly, an interaction library must be large enough so that at least one of its members will have a structure that gives it affinity for the target analyte. Although it is difficult to gauge the required absolute size of an interaction library, nature provides a hint with the immune response: a diversity of 10^7-10^8 different antibodies provides at least one combination with sufficient affinity to interact with most potential antigens faced by an organism. Published in vitro selection techniques have also shown that a library size of 10^7 to 10^8 is sufficient to find structures with affinity for the target. Thus, in a preferred embodiment, at least 10^6, preferably at least 10^7, more preferably at least 10^8 and most preferably at least 10^9 different bioactive agents are simultaneously analyzed in the subject methods. Preferred methods maximize library size and diversity.

In a preferred embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

In a preferred embodiment, the bioactive agents are nucleic acids (generally called "probe nucleic acids" or "candidate probes" herein). By "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, as outlined below, nucleic acid analogs are included that may have alternate backbones, comprising, for example, phosphoramide (Beaucage, *et al.*, Tetrahedron, 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem., 35:3800 (1970); Sprinzl, *et al.*, Eur. J. Biochem., 81:579 (1977); Letsinger, *et al.*, Nucl. Acids Res., 14:3487 (1986); Sawai, *et al.*, Chem. Lett., 805 (1984), Letsinger, *et al.*, J. Am. Chem. Soc., 110:4470 (1988); and Pauwels, *et al.*, Chemica Scripta, 26:141 (1986)), phosphorothioate (Mag, *et al.*, Nucleic Acids Res., 19:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu, *et al.*, J. Am. Chem. Soc., 111:2321 (1989)), O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, J. Am. Chem. Soc., 114:1895 (1992); Meier, *et al.*, Chem. Int. Ed. Engl., 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson, *et al.*, Nature, 380:207 (1996), all of which are incorporated by reference)). Other analog nucleic acids include those with positive backbones (Denpcy, *et al.*, Proc. Natl. Acad. Sci. USA, 92:6097 (1995)); non-ionic backbones (U.S. Patent Nos. 5,386,023; 5,637,684; 5,602,240; 5,216,141; and 4,469,863; Kiedrowshi, *et al.*, Angew. Chem. Intl. Ed. English, 30:423 (1991); Letsinger, *et al.*, J. Am. Chem. Soc., 110:4470 (1988); Letsinger, *et al.*, Nucleosides & Nucleotides, 13:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker, *et al.*, Bioorganic & Medicinal Chem. Lett., 4:395 (1994); Jeffs, *et al.*, J. Biomolecular NMR, 34:17 (1994); Tetrahedron Lett., 37:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins, *et al.*, Chem. Soc. Rev., (1995) pp. 169-176). Several nucleic acid analogs are described in Rawls, C & E News, June 2, 1997, page 35. All of these references are hereby expressly incorporated by reference. These modifications of the ribose-phosphate backbone may be done to facilitate the addition of additional moieties such as labels, or to increase the stability and half-life of such molecules in physiological environments; for example, PNA is particularly preferred. In addition, mixtures of naturally occurring nucleic acids and analogs can be made. Alternatively,

mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made. The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthanine, hypoxanthanine, isocytosine, isoguanine, and base analogs such as nitropyrrole and nitroindole, etc.

As described above generally for proteins, nucleic acid bioactive agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of procaryotic or eukaryotic genomes may be used as is outlined above for proteins.

In general, probes of the present invention are designed to be complementary to a target sequence (either the target analyte sequence of the sample or to other probe sequences, as is described herein), such that hybridization of the target and the probes of the present invention occurs. This complementarity need not be perfect; there may be any number of base pair mismatches that will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under the selected reaction conditions. High stringency conditions are known in the art; see for example Maniatis et al., Molecular Cloning: A Laboratory Manual, 2d Edition, 1989, and Short Protocols in Molecular Biology, ed. Ausubel, et al., both of which are hereby incorporated by reference. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Techniques in Biochemistry and Molecular Biology–Hybridization with Nucleic Acid Probes, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. In another embodiment, less stringent hybridization conditions are used; for example, moderate or low stringency conditions may be used, as are known in the art; see Maniatis and Ausubel, supra, and Tijssen, supra.

The term 'target sequence" or grammatical equivalents herein means a nucleic acid sequence on a single strand of nucleic acid. The target sequence may be a portion of a gene, a regulatory sequence, genomic DNA, cDNA, RNA including mRNA and rRNA, or others. It may be any length, with the understanding that longer sequences are more specific. As will be appreciated by those in the art, the complementary target sequence may take many forms. For example, it may be contained within a larger nucleic acid sequence, i.e. all or part of a gene or mRNA, a restriction fragment of a plasmid or genomic DNA, among others. As is outlined more fully below, probes are made to hybridize to target sequences to determine the presence or absence of the target sequence in a sample. Generally speaking, this term will be understood by those skilled in the art.

In a preferred embodiment, the bioactive agents are organic chemical moieties, a wide variety of which are available in the literature.

In a preferred embodiment, each bead comprises a single type of bioactive agent, although a plurality of individual bioactive agents are preferably attached to each bead. Similarly, preferred embodiments utilize more than one microsphere containing a unique bioactive agent; that is, there is redundancy built into the system by the use of subpopulations of microspheres, each microsphere in the subpopulation containing the same bioactive agent.

As will be appreciated by those in the art, the bioactive agents may either be synthesized directly on the beads, or they may be made and then attached after synthesis. In a preferred embodiment, linkers are used to attach the bioactive agents to the beads, to allow both good attachment, sufficient flexibility to allow good interaction with the target molecule, and to avoid undesirable binding reactions.

In a preferred embodiment, the bioactive agents are synthesized directly on the beads. As is known in the art, many classes of chemical compounds are currently synthesized on solid supports, such as peptides, organic moieties, and nucleic acids. It is a relatively straightforward matter to adjust the current synthetic techniques to use beads.

In a preferred embodiment, the bioactive agents are synthesized first, and then covalently attached to the beads. As will be appreciated by those in the art, this will be done depending on the composition of the bioactive agents and the beads. The functionalization of solid support surfaces such as certain polymers with chemically reactive groups such as thiols, amines, carboxyls, etc. is generally known in the art. Accordingly, "blank" microspheres may be used that have surface chemistries that facilitate the attachment of the desired functionality by the user. Some examples of these surface chemistries for blank microspheres include, but are not limited to, amino groups including aliphatic and aromatic amines, carboxylic acids, aldehydes, amides, chloromethyl groups, hydrazide, hydroxyl groups, sulfonates and sulfates.

These functional groups can be used to add any number of different candidate agents to the beads, generally using known chemistries. For example, candidate agents containing carbohydrates may be attached to an amino-functionalized support; the aldehyde of the carbohydrate is made using standard techniques, and then the aldehyde is reacted with an amino group on the surface. In an alternative embodiment, a sulfhydryl linker may be used. There are a number of sulfhydryl reactive linkers known in the art such as SPDP, maleimides, α-haloacetyls, and pyridyl disulfides (see for example the 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference) which can be used to attach cysteine containing proteinaceous agents to the support. Alternatively, an amino group on the candidate agent may be used for attachment to an amino group on the surface. For example, a large number of stable bifunctional groups are well known in the art, including homobifunctional and heterobifunctional linkers (see Pierce Catalog and Handbook, pages 155-200). In an additional embodiment, carboxyl groups (either from the surface or from the candidate agent) may be derivatized using well known linkers (see the Pierce catalog). For example, carbodiimides activate carboxyl groups for attack by good nucleophiles such as amines (see Torchilin et al., Critical Rev. Therapeutic Drug Carrier Systems, 7(4):275-308 (1991), expressly incorporated herein). Proteinaceous candidate agents may also be attached using other techniques known in the art, for example for the attachment of antibodies to polymers; see Slinkin et al., Bioconj. Chem. 2:342-348 (1991); Torchilin et al., *supra*; Trubetskoy et al., Bioconj. Chem. 3:323-327 (1992); King et al., Cancer Res. 54:6176-6185 (1994); and Wilbur et al., Bioconjugate Chem. 5:220-235 (1994), all of which are hereby expressly incorporated by reference). It should be understood that the candidate agents may be attached in a variety of ways, including those listed above. What is important is that manner of attachment does not significantly alter the functionality of the candidate agent; that is, the candidate agent should be attached in such a flexible manner as to allow its interaction with a target.

Specific techniques for immobilizing enzymes on microspheres are known in the prior art. In one case, NH_2 surface chemistry microspheres are used. Surface activation is achieved with a 2.5% glutaraldehyde in phosphate buffered saline (10 mM) providing a pH of 6.9. (138 mM NaCl, 2.7 mM, KCl). This is stirred on a stir bed for approximately 2 hours at room temperature. The microspheres are then rinsed with ultrapure water plus 0.01% tween 20 (surfactant) -0.02%, and rinsed again with a pH 7.7 PBS plus 0.01% tween 20. Finally, the enzyme is added to the solution, preferably after being prefiltered using a 0.45μm amicon micropure filter.

In some embodiments, the microspheres may additionally comprise identifier binding ligands for use in certain decoding systems. By "identifier binding ligands" or "IBLs" herein is meant a compound that will specifically bind a corresponding decoder binding ligand (DBL) to facilitate the elucidation of the identity of the bioactive agent attached to the bead. That is, the IBL and the corresponding DBL form a binding partner pair. By "specifically bind" herein is meant that the IBL binds its DBL with specificity sufficient to differentiate between the corresponding DBL and other DBLs (that is, DBLs for other

IBLs), or other components or contaminants of the system. The binding should be sufficient to remain bound under the conditions of the decoding step, including wash steps to remove non-specific binding. In some embodiments, for example when the IBLs and corresponding DBLs are proteins or nucleic acids, the dissociation constants of the IBL to its DBL will be less than about 10^{-4}-10^{-6} M^{-1}, with less than about 10^{-5} to 10^{-9} M^{-1} being preferred and less than about 10^{-7} -10^{-9} M^{-1} being particularly preferred.

IBL-DBL binding pairs are known or can be readily found using known techniques. For example, when the IBL is a protein, the DBLs include proteins (particularly including antibodies or fragments thereof (FAbs, etc.)) or small molecules, or vice versa (the IBL is an antibody and the DBL is a protein). Metal ion- metal ion ligands or chelators pairs are also useful. Antigen-antibody pairs, enzymes and substrates or inhibitors, other protein-protein interacting pairs, receptor-ligands, complementary nucleic acids, and carbohydrates and their binding partners are also suitable binding pairs. Nucleic acid - nucleic acid binding proteins pairs are also useful. Similarly, as is generally described in U.S. Patents 5,270,163, 5,475,096, 5,567,588, 5,595,877, 5,637,459, 5,683,867,5,705,337, and related patents, hereby incorporated by reference, nucleic acid "aptomers" can be developed for binding to virtually any target; such a aptomer-target pair can be used as the IBL-DBL pair. Similarly, there is a wide body of literature relating to the development of binding pairs based on combinatorial chemistry methods.

In a preferred embodiment, the IBL is a molecule whose color or luminescence properties change in the presence of a selectively-binding DBL. For example, the IBL may be a fluorescent pH indicator whose emission intensity changes with pH. Similarly, the IBL may be a fluorescent ion indicator, whose emission properties change with ion concentration.

Alternatively, the IBL is a molecule whose color or luminescence properties change in the presence of various solvents. For example, the IBL may be a fluorescent molecule such as an ethidium salt whose fluorescence intensity increases in hydrophobic environments. Similarly, the IBL may be a derivative of fluorescein whose color changes between aqueous and nonpolar solvents.

In one embodiment, the DBL may be attached to a bead, i.e. a "decoder bead", that may carry a label such as a fluorophore.

In a preferred embodiment, the IBL-DBL pair comprise substantially complementary single-stranded nucleic acids. In this embodiment, the binding ligands can be referred to as "identifier probes" and "decoder probes". Generally, the identifier and decoder probes range from about 4 basepairs in length to about 1000, with from about 6 to about 100 being preferred, and from about 8 to about 40 being particularly preferred. What is important is that the probes are long enough to be specific, i.e. to

distinguish between different IBL-DBL pairs, yet short enough to allow both a) dissociation, if necessary, under suitable experimental conditions, and b) efficient hybridization.

In a preferred embodiment, as is more fully outlined below, the IBLs do not bind to DBLs. Rather, the IBLs are used as identifier moieties ("IMs") that are identified directly, for example through the use of mass spectroscopy.

In a preferred embodiment, the microspheres comprise an optical signature that can be used to identify the attached bioactive agent, as is generally outlined in U.S.S.N.s 08/818,199 and 09/151,877, both of which are hereby incorporated by reference. That is, each subpopulation of microspheres comprise a unique optical signature or optical tag that can be used to identify the unique bioactive agent of that subpopulation of microspheres; a bead comprising the unique optical signature may be distinguished from beads at other locations with different optical signatures. As is outlined herein, each bioactive agent will have an associated unique optical signature such that any microspheres comprising that bioactive agent will be identifiable on the basis of the signature. As is more fully outlined below, it is possible to reuse or duplicate optical signatures within an array, for example, when another level of identification is used, for example when beads of different sizes are used, or when the array is loaded sequentially with different batches of beads.

In a preferred embodiment, the optical signature is generally a mixture of reporter dyes, preferably fluorescent. By varying both the composition of the mixture (i.e. the ratio of one dye to another) and the concentration of the dye (leading to differences in signal intensity), matrices of unique tags may be generated. This may be done by covalently attaching the dyes to the surface of the beads, or alternatively, by entrapping the dye within the bead. The dyes may be chromophores or phosphors but are preferably fluorescent dyes, which due to their strong signals provide a good signal-to-noise ratio for decoding. Suitable dyes for use in the invention include, but are not limited to, fluorescent lanthanide complexes, including those of Europium and Terbium, fluorescein, rhodamine, tetramethylrhodamine, eosin, erythrosin, coumarin, methyl-coumarins, pyrene, Malacite green, stilbene, Lucifer Yellow, Cascade Blue™, Texas Red, and others described in the 1989-1991 Molecular Probes Handbook by Richard P. Haugland, hereby expressly incorporated by reference.

In a preferred embodiment, the encoding can be accomplished in a ratio of at least two dyes, although more encoding dimensions may be added in the size of the beads, for example. In addition, the labels are distinguishable from one another; thus two different labels may comprise different molecules (i.e. two different fluors) or, alternatively, one label at two different concentrations or intensity.

In a preferred embodiment, the dyes are covalently attached to the surface of the beads. This may be done as is generally outlined for the attachment of the bioactive agents, using functional groups on the

surface of the beads. As will be appreciated by those in the art, these attachments are done to minimize the effect on the dye.

In a preferred embodiment, the dyes are non-covalently associated with the beads, generally by entrapping the dyes in the bead matrix or pores of the beads. Fluorescent dyes are generally preferred because the strength of the fluorescent signal provides a better signal-to-noise ratio when decoding. Additionally, encoding in the ratios of the two or more dyes, rather than single dye concentrations, is preferred since it provides insensitivity to the intensity of light used to interrogate the reporter dye's signature and detector sensitivity.

In one embodiment, the dyes are added to the bioactive agent, rather than the beads, although this is generally not preferred.

In one embodiment, the microspheres do not contain an optical signature.

In one embodiment the microspheres contain a signal transducer that indicates and/or quantifies the recognition event. By "signal transducer" or "signal transducer element" herein is meant a molecule whose detectable signal is altered upon association of a target analyte with a bioactive agent. In some embodiments the bioactive agent is itself the signal transducer, however, this is not preferred.

In a preferred embodiment, the microspheres contain a label or signal transducer that transduces a signal upon association of a bioactive agent with a target analyte. In a preferred embodiment the signal transducer produces an optical signal that upon binding of the target analyte to the bioactive agent produces and/or changes the optical properties of the signal transducer. Examples of signal transducers include but are not limited to nucleotides (i.e. DNA or RNA) intercalators, fluorophores and the like. In addition, signal transducers are characterized by detecting disturbance of a fluorescence resonance energy transfer system, disturbance of an electron transfer system, disturbance of a micro-environmentally sensitive fluorophore, inhibition of enzyme reaction upon binding to a target analyte, or fluorophores that change their fluorescent properties upon recognition of analyte.

In one embodiment, the signal transducer is a fluorophore. The fluorophore is attached to the beads either directly or indirectly. Upon binding or association of the target analyte, the emission of the fluorophore changes such that the change is an indication of the binding of the target analyte to the bead and/or bioactive agent (Figure 9).

Alternatively, the bioactive agent is attached to the microsphere either directly, or indirectly, for example by binding the signal transducer on the bead. In this embodiment, upon binding of the target analyte to the bioactive agent, the emission of the fluorophore changes such that the change is an indication of the binding of the target analyte to the bioactive agent (Figure 10).

In an alternative embodiment, the bioactive agent is indirectly attached or bound to a signal transducer through a generic or universal tag (see Figure 11). In this embodiment, the bioactive agent includes a binding moiety that binds the universal tag and also includes a recognition moiety for binding the target analyte (Figure 11). Upon hybridization of the target analyte (that may or may not contain a label), the emission or signal of the signal tranduces changes such that detection fo the change is an indication of the binding of the target analyte to the bioactive agent.

Accordingly, In this embodiment, the array contains microspheres that contain optical transduction mechanisms i.e. signal transducer elements, that allow direct sensing of the analyte without pre-labeling the target analyte or adding auxiliary reporter agents. That is, the invention provides arrays for direct sensing of unlabeled analytes.

The detection method will typically also incorporate some signal processing to determine whether the signal at a particular matrix position is a true positive or may be a spurious signal. For example, a signal from a region which has actual positive signal may tend to spread over and provide a positive signal in an adjacent region which actually should not have one. This may occur, e.g., where the scanning system is not properly discriminating with sufficiently high resolution in its pixel density to separate the two regions. Thus, the signal over the spatial region may be evaluated pixel by pixel to determine the locations and the actual extent of positive signal. A true positive signal should, in theory, show a uniform signal at each pixel location. Thus, processing by plotting number of pixels with actual signal intensity should have a clearly uniform signal intensity. Regions where the signal intensities show a fairly wide dispersion, may be particularly suspect and the scanning system may be programmed to more carefully scan those positions. Doping is a technique designed to reduce the overall capacitance of the output sense node. The effect is to increase the conversion gain of the output amplifier and reduce the noise.

In a preferred embodiment, the surface of the substrate is modified to contain wells, i.e. depressions in the surface of the substrate. This may be done as is generally known in the art using a variety of techniques, including, but not limited to, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those in the art, the technique used will depend on the composition and shape of the substrate. When the first substrate comprises both the assay locations and the individual arrays, a preferred method utilizes molding techniques that form the bead wells in the bottom of the assay wells in a microtiter plate. Similarly, a preferred embodiment utilizes a molded second substrate, comprising "fingers" or projections in an array format, and each finger comprises bead wells.

In a preferred embodiment, the sites or wells are separated with spaces between each other. Prior art suggests that the relationship between the test sites and the pixel is that multiple pixels are necessary to resolve the image of a signal at a particular site on an array, e.g. a particular bead. However, the

pixels used to resolve the signal duplicate information, and ultimately provide extraneous information. One array site such as a well or bead imaged in one pixel is a much more efficient use of the finite number of pixels. Sub-pixel resolution, or below resolution of a pixel will allow for imaging of more samples per assay.

As is appreciated by those skilled in the relevant art, bead spacing is determined by calculating the distance between centers. Varying the spacing between sites results in the formation of arrays of high density, medium density or lower density. High density arrays are characterized as having sites separated by less than about 5 to 15 μm. Medium density arrays have sites separated by about 15 to 30 μm, while low density arrays have sites separated by greater than 30 μm. Generally, the sites are separated by less than 100 μm; preferably less than 50 μm and most preferably less than 15-20 μm. A particular advantage of spacing wells apart is that commercial scanners can be used to analyze the arrays. That is, the resolution of scanners varies and arrays can be formed that allow for detection on high or low resolution scanners. For high density arrays, high resolution scanners (< 5μm) can be employed. These scanners effectively analyze arrays with close spacing (< 15 μm) between features, i.e. beads. For lower resolution scanners (> 5 μm), increased bead spacing, i.e. > 10 μm can be utilized, with from 15 to 20 μm being preferred. In both cases, various software packages are used, such as but not limited to, GENEPIX software package by AXON instruments or others that are provided with conventional fluorescent microscope scanning equipment. In a preferred embodiment, the software employs contrast-based or other image processing algorithms to resolve the beads and extract signal intensity information (see also USSN 09/651,181, filed August 30, 2000 and PCT/US00/23830, filed August 30, 2000, both of which are expressly incorporated herein by reference).

While in the above described embodiment the spacing between features is accomplished by physically altering the spacing of the sites on the substrate, in an alternative embodiment, when beads in bead wells form the array, density is modulated by adding to a population of beads comprising bioactive agents, a population of beads that do not comprise a bioactive agent. That is, a population of beads, with no bioactive agent, and in some embodiments no detectable signal or label, is added to at least one population of beads that does comprise a bioactive agent. The beads lacking a bioactive agent, i.e. "blank beads", dilute the concentration of beads with a bioactive agent. When applied to or distributed on a substrate, this results in increased spacing between beads with bioactive agents. That is, in the absence of blank beads, beads with bioactive agents will fill substantially all of the wells on a substrate at an average density of not more than one bead per well. When the spacing of wells is close, only high resolution scanners effectively analyze the array. However, upon the addition of a population of blank beads, blank beads will be distributed with the beads that have bioactive agents thereby increasing the distance between beads with bioactive agents. Thus, in a preferred embodiment, the distance between centers of beads with bioactive agents is at least 5 μm; more preferably between 10 to 50 μm; and most preferably between 15 to 25 μm.

In one embodiment, the ratio between beads with bioactive agents and blank beads is adjusted to achieve proper density of beads with bioactive agents on the array. The ratio depends on the desired spacing between beads, which is generally based on the resolution of the scanner. While in some embodiments it may only be necessary to include a small number of blank beads, i.e. the ratio is about 10:1 or greater, in other embodiments, the ratio is at least 1:36 or more, or 1:100.

In an alternative embodiment, the array comprises a first population of beads with a first bioactive agent and a second population of beads with a second population of bioactive agents. When modulating the spacing of beads on an array so that conventional scanners can be used, it is useful in this embodiment for each population of beads to be labeled or tagged with different tags. The tags are preferably detectable in distinct channels. As such, only one population of beads is analyzed at a time. Accordingly, the beads that are not being analyzed serve as spacer beads although they do contain a bioactive agent and can be analyzed in a different channel. As such, the spacing of the beads from each population will be adequately spaced for analysis, while the the number of beads to be analyzed is increased relative the above-described assay that uses blank beads. That is, when analyzing the first population of beads in a first channel, which does not detect the second population of beads, the second population of beads serve as spacing beads or blank beads. The second population of beads serves to increase the spacing between the first population of beads. In turn, when analyzing the second population of beads in a second channel, the first population of beads serves as spacing or "blank" beads that separate the second population of beads.

Once made, the array finds use in a variety of assays, including analyzing complex mixtures. In a preferred embodiment, the mixture is not labeled. In this embodiment, the change in signal transduction properties of the signal transducer on the microsphere(s) provides the signal or indicates the presence of the target analyte in the mixture.

Additional advantages of the improved array include a reduction in sample preparation time and expense, no labeling of the sample, increased efficiency of sample preparation and reduction in signal background as a result of unreacted reporter groups. In addition, generic or universal platforms can be designed. That is, arrays can be designed for the detection of a variety of target analytes; however, there is no need for alternative sample preparation or labeling.

In a preferred embodiment, the array finds use in a kit for the detection of target analytes. The kit includes microspheres that contain a bioactive agent and a signal transducer. In addition, the kit contains a substrate with a surface with discrete sites. The microspheres are distributed on the substrate either prior to, simultaneously with or subsequent to contacting of the beads with the sample. In one embodiment, the signal transducer is attached to the bioactive agent.

In a preferred embodiment, the present invention does not rely solely on the use of optical properties to decode the arrays. However, as will be appreciated by those in the art, it is possible in some embodiments to utilize optical signatures as an additional coding method, in conjunction with the present system. Thus, for example, as is more fully outlined below, the size of the array may be effectively increased while using a single set of decoding moieties in several ways, one of which is the use of optical signatures one some beads. Thus, for example, using one "set" of decoding molecules, the use of two populations of beads, one with an optical signature and one without, allows the effective doubling of the array size. The use of multiple optical signatures similarly increases the possible size of the array.

In a preferred embodiment, each subpopulation of beads comprises a plurality of different identifier binding ligands ("IBLs"). By using a plurality of different IBLs to encode each bioactive agent, the number of possible unique codes is substantially increased. That is, by using one unique IBL per bioactive agent, the size of the array will be the number of unique IBLs (assuming no "reuse" occurs, as outlined below). However, by using a plurality of different IBLs per bead, n, the size of the array can be increased to 2^n, when the presence or absence of each IBL is used as the indicator. For example, the assignment of 10 IBLs per bead generates a 10 bit binary code, where each bit can be designated as "1" (IBL is present) or "0" (IBL is absent). A 10 bit binary code has 2^{10} possible variants However, as is more fully discussed below, the size of the array may be further increased if another parameter is included such as concentration or intensity; thus for example, if two different concentrations of the IBL are used, then the array size increases as 3^n. Thus, in this embodiment, each individual bioactive agent in the array is assigned a combination of IBLs, which can be added to the beads prior to the addition of the bioactive agent, after, or during the synthesis of the bioactive agent, i.e. simultaneous addition of IBLs and bioactive agent components.

Alternatively, when the bioactive agent is a polymer of different residues, i.e. when the bioactive agent is a protein or nucleic acid, the combination of different IBLs can be used to elucidate the sequence of the protein or nucleic acid.

Thus, for example, using two different IBLs (IBL1 and IBL2), the first position of a nucleic acid can be elucidated: for example, adenosine can be represented by the presence of both IBL1 and IBL2; thymidine can be represented by the presence of IBL1 but not IBL2, cytosine can be represented by the presence of IBL2 but not IBL1, and guanosine can be represented by the absence of both. The second position of the nucleic acid can be done in a similar manner using IBL3 and IBL4; thus, the presence of IBL1, IBL2, IBL3 and IBL4 gives a sequence of AA; IBL1, IBL2, and IBL3 shows the sequence AT; IBL1, IBL3 and IBL4 gives the sequence TA, etc. The third position utilizes IBL5 and IBL6, etc. In this way, the use of 20 different identifiers can yield a unique code for every possible 10-mer.

The system is similar for proteins but requires a larger number of different IBLs to identify each position, depending on the allowed diversity at each position. Thus for example, if every amino acid is allowed at every position, five different IBLs are required for each position. However, as outlined above, for example when using random peptides as the bioactive agents, there may be bias built into the system; not all amino acids may be present at all positions, and some positions may be preset; accordingly, it may be possible to utilize four different IBLs for each amino acid.

In this way, a sort of "bar code" for each sequence can be constructed; the presence or absence of each distinct IBL will allow the identification of each bioactive agent.

In addition, the use of different concentrations or densities of IBLs allows a "reuse" of sorts. If, for example, the bead comprising a first agent has a 1X concentration of IBL, and a second bead comprising a second agent has a 10X concentration of IBL, using saturating concentrations of the corresponding labelled DBL allows the user to distinguish between the two beads.

In a preferred embodiment, the compositions of the invention further comprise at least one fiducial. By "fiducial" or "marker" or "registration point" herein is meant a physical reference feature or characteristic that allows precise comparisons of sequential data images of an array. The use of fiducials is useful for a variety of reasons. In general, the assays involve monitoring of objects, i.e. bioactive agents, located at spatially distinct locations (features) over the course of several data image frames taken over time. Any shifting that occurs from frame to frame complicates the analysis of the agents. By incorporating permanent fiducials into the assay structure, each data image can be aligned, either manually or automatically, to allow accurate comparison of the images, and control for translation (i.e. a shift in an X-Y direction) and/or rotation as well as reduction or enlargement of the image. In addition, when fluorescence based assays are used (either for decoding or analyte assaying or both), in any given image, a particular region or feature may or may not emit fluorescence, depending on the label characteristics and the wavelength being interrogated, or the presence or absence of an analyte or DBL, etc. The presence of fluorescence is detected as a positive change in feature intensity with respect to the background intensity, which is then used to draw a software "segment" over the core. In situations where the core is dark, i.e. no fluorescence is detected at that particular feature, it is difficult to accurately draw the segment over the core.

Accordingly, in a preferred embodiment, at least one fiducial is incorporated into the array. In a preferred embodiment, a plurality of fiducials are used, with the ideal number depending on the size of the array (i.e. features per fiducial), the density of the array, the shape of the array, the irregularity of the array, etc. In general, at least three non-linear fiducials are used; that is, three fiducials that define a plane (i.e. are not in a line) are used. In addition, it is preferred to have at least one of the fiducials be either on or close to the periphery of the array. Fiducials are described in more detail in U.S.S.N. 09/500,555, and U.S. Application entitled Automated Information Processing in Randomly Ordered

Arrays, filed August 9, 2000 (no serial number received), both of which are hereby expressly incorporated by reference.

Once the microspheres comprising the candidate agents and the unique tags are generated, they are added to the substrate to form an array. In general, the methods of making the arrays and of decoding the arrays is done to maximize the number of different candidate agents that can be uniquely encoded. The compositions of the invention may be made in a variety of ways. In general, the arrays are made by adding a solution or slurry comprising the beads to a surface containing the sites for attachment of the beads. This may be done in a variety of buffers, including aqueous and organic solvents, and mixtures. The solvent can evaporate, and excess beads removed.

In some embodiments, for example when chemical attachment is done, it is possible to attach the beads in a non-random or ordered way. For example, using photoactivatible attachment linkers or photoactivatible adhesives or masks, selected sites on the array may be sequentially rendered suitable for attachment, such that defined populations of beads are laid down.

The arrays of the present invention are constructed such that information about the identity of the candidate agent is built into the array, such that the random deposition of the beads in the fiber wells can be "decoded" to allow identification of the candidate agent at all positions. This may be done in a variety of ways, and either before, during or after the use of the array to detect target molecules.

Thus, after the array is made, it is "decoded" in order to identify the location of one or more of the bioactive agents, i.e. each subpopulation of beads, on the substrate surface. In general, both decoding and the experimental assay to determine the presence or absence of a target analyte, both of which are described below, requires the comparison of sequential data images to determine the differences between two data images. In general, this is done by taking a first or initial data image, using the fiducial to create a registered first data image, subjecting the array to decoding conditions and taking a second data image. The same fiducial is used to create a registered second data image, and then the two registered images can be compared. In this context, a "data image" includes a primary data image or a reduction of the image; for example, the image may be reduced to a set of X-Y coordinates with corresponding intensity values.

In a preferred embodiment, this is done using a computer system comprising a processor and a computer readable memory. The computer readable memory comprises an acquisition module that comprises computer code that can receive a data image from a random array and a registration module comprising computer code that can register the data image using at least one fiducial, including a fiducial template, to generate a registered data image. This registered data image can then be stored in a storage module as needed. This same computer code, or different code, if required, can be used to receive additional data images and generate additional registered data

images, which also can be stored. The computer readable memory further comprises a comparison module comprising computer code that can compare the registered data images to determine the differences between them, to allow both decoding of the array and target analyte detection. That is, when decoding is done, the comparison of at least two registered data images allows the identification of the location of at least two unique bioactive agents on the array.

As a preliminary matter, prior to decoding, a filtering step or preprocessing step is performed (although in some embodiments this step is performed during or after decoding). That is, in some embodiments, filtering or preprocessing is performed on the array. In one embodiment, the preprocessing serves to identify the array elements that contain no beads or contain faulty beads. That is, as described in detail below, preprocessing identifies the array locations that do not contain a detectable signal or contain a signal that is not similar to a reference signal or signals. These locations are defined as containing no beads or containing faulty beads and thus can be dropped from the analysis, giving a higher confidence level for the remaining sites.

In an alternative embodiment, filtering or preprocessing serves to categorize the beads into subpopulations; i.e. beads with similar characteristics, such as color, may form a subpopulation. The subpopulations may be used for additional data processing such as signal summing, statistical analyses or comparison. This may also serve to allow outliers within a subpopulation to be identified and discarded, resulting in higher confidence levels, higher signals and lower background.

Generally, the preprocessing is performed by analyzing or detecting a signal obtained from at least one of the array locations and determining whether the array location contains a bead. If the array location does contain a bead with a detectable signal, the bead may be further categorized into subpopulations containing similar signals, although this is not required. If it is determined that the assay location does not contain a bead, the assay location may be disregarded during additional analyses.

In one embodiment, the signal obtained from array location is an optical signature of the bead. That is, as described herein, beads or microspheres can be labeled directly or indirectly with dyes or fluorophores. Any of the optical signatures as described herein find use with preprocessing of the array image.

In addition, the signal(s) from the clustered beads can then be analyzed by summing the signals or other statistical analyses. Signal summing and statistical analyses are more thoroughly described in USSNs 08/6944,850 and 09/287,573, 06/238,866 and PCT/US98/21193, PCT/US01/31581 and PCT US00/09183, all of which are expressly incorporated herein by reference in their entirety.

Thus, using the systems described herein, a random array is decoded as is generally described in USSNs 60/090,473, 09/189,543 and 09/344,526 and PCT/US99/14387, all of which are expressly incorporated herein by reference in their entirety. In a preferred embodiment, a selective decoding system is used. In this case, only those microspheres exhibiting a change in the optical signal as a result of the binding of a target analyte are decoded. This is commonly done when the number of "hits", i.e. the number of sites to decode, is generally low. That is, the array is first scanned under experimental conditions in the absence of the target analytes. The sample containing the target analytes is added, and only those locations exhibiting a change in the optical signal are decoded. For example, the beads at either the positive or negative signal locations may be either selectively tagged or released from the array (for example through the use of photocleavable linkers), and subsequently sorted or enriched in a fluorescence-activated cell sorter (FACS). That is, either all the negative beads are released, and then the positive beads are either released or analyzed in situ, or alternatively all the positives are released and analyzed. Alternatively, the labels may comprise halogenated aromatic compounds, and detection of the label is done using for example gas chromatography, chemical tags, isotopic tags mass spectral tags.

In addition using the systems described herein, composite arrays are generally described in USSNs 09/782,588, 09/931,271, 60/181,163, and PCT/US01/04504, all of which are expressly incorporated herein by reference in their entirety.

In a preferred embodiment, atomic force microscopy (AFM) is used to decode the array. In this embodiment, an AFM tip, comprising a DBL, is positioned at the site to be decoded, that comprises an IBL. The force of interaction between the IBL/DBL is measured using AFM. In addition, since AFM has atomic resolution, a variety of other physical characteristics, including physical size and shape can be used for decoding. For example, different "shaped" molecules could be used as IBLs; in this embodiment, the AFM tip can comprise a DBL or just a moiety that can detect different surfaces. In addition, AFM could be used as "nanotweezers" to deliver or recover beads to and from specific locations on the array.

As will be appreciated by those in the art, this may also be done in systems where the array is not decoded; i.e. there need not ever be a correlation of bead composition with location. In this embodiment, the beads are loaded on the array, and the assay is run. The "positives", i.e. those beads displaying a change in the optical signal as is more fully outlined below, are then "marked" to distinguish or separate them from the "negative" beads. This can be done in several ways, preferably using fiber optic arrays. In a preferred embodiment, each bead contains a fluorescent dye. After the assay and the identification of the "positives" or "active beads", light is shown down either only the positive fibers or only the negative fibers, generally in the presence of a light-activated reagent (typically dissolved oxygen). In the former case, all the active beads are photobleached. Thus, upon non-selective release of all the beads with subsequent sorting, for example using a fluorescence

activated cell sorter (FACS) machine, the non-fluorescent active beads can be sorted from the fluorescent negative beads. Alternatively, when light is shown down the negative fibers, all the negatives are non-fluorescent and the positives are fluorescent, and sorting can proceed. The characterization of the attached bioactive agent may be done directly, for example using mass spectroscopy.

In one embodiment, when chiral molecules, such as DNA, are to be detected on the array, the samples or target molecules need not be labeled. That is, the molecules are detected on arrays without being labeled. The chiral molecule may be a bioactive agent on a microsphere, or alternatively, it may be the target analyte. In one embodiment the chiral molecule is attached directly to the substrate. In an alternative embodiment the molecule is indirectly attached to the substrate, for example, via an intermediate moiety such as a microsphere that is distributed on the substrate. In both of these embodiments, plane-or circularly polarized light is shined or illuminated onto an array. The absorption or emission properties of the array are determined. Those features of the array that contain the chiral molecule will rotate the light. Those features of the array to which no chiral molecule is bound will not display rotated light. Accordingly, the invention provides a method for distinguishing array features that contain a chiral molecule from those that do not.

Alternatively, the identification may occur through the use of identifier moieties ("IMs"), which are similar to IBLs but need not necessarily bind to DBLs. That is, rather than elucidate the structure of the bioactive agent directly, the composition of the IMs may serve as the identifier. Thus, for example, a specific combination of IMs can serve to code the bead, and be used to identify the agent on the bead upon release from the bead followed by subsequent analysis, for example using a gas chromatograph or mass spectroscope.

Alternatively, rather than having each bead contain a fluorescent dye, each bead comprises a non-fluorescent precursor to a fluorescent dye. For example, using photocleavable protecting groups, such as certain ortho-nitrobenzyl groups, on a fluorescent molecule, photoactivation of the fluorochrome can be done. After the assay, light is shown down again either the "positive" or the "negative" fibers, to distinguish these populations. The illuminated precursors are then chemically converted to a fluorescent dye. All the beads are then released from the array, with sorting, to form populations of fluorescent and non-fluorescent beads (either the positives and the negatives or vice versa).

In an alternate preferred embodiment, the sites of attachment of the beads (for example the wells) include a photopolymerizable reagent, or the photopolymerizable agent is added to the assembled array. After the test assay is run, light is shown down again either the "positive" or the "negative" fibers, to distinguish these populations. As a result of the irradiation, either all the positives or all the

negatives are polymerized and trapped or bound to the sites, while the other population of beads can be released from the array.

In a preferred embodiment, the location of every bioactive agent is determined using decoder binding ligands (DBLs). As outlined above, DBLs are binding ligands that will either bind to identifier binding ligands, if present, or to the bioactive agents themselves, preferably when the bioactive agent is a nucleic acid or protein.

In a preferred embodiment, as outlined above, the DBL binds to the IBL.

In a preferred embodiment, the bioactive agents are single-stranded nucleic acids and the DBL is a substantially complementary single-stranded nucleic acid that binds (hybridizes) to the bioactive agent, termed a decoder probe herein. A decoder probe that is substantially complementary to each candidate probe is made and used to decode the array. In this embodiment, the candidate probes and the decoder probes should be of sufficient length (and the decoding step run under suitable conditions) to allow specificity; i.e. each candidate probe binds to its corresponding decoder probe with sufficient specificity to allow the distinction of each candidate probe.

In a preferred embodiment, the DBLs are either directly or indirectly labeled. By "labeled" herein is meant that a compound has at least one element, isotope or chemical compound attached to enable the detection of the compound. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) magnetic, electrical, thermal; and c) colored or luminescent dyes; although labels include enzymes and particles such as magnetic particles as well. Preferred labels include luminescent labels. In a preferred embodiment, the DBL is directly labeled, that is, the DBL comprises a label. In an alternate embodiment, the DBL is indirectly labeled; that is, a labeling binding ligand (LBL) that will bind to the DBL is used. In this embodiment, the labeling binding ligand-DBL pair can be as described above for IBL-DBL pairs.

Accordingly, the identification of the location of the individual beads (or subpopulations of beads) is done using one or more decoding steps comprising a binding between the labeled DBL and either the IBL or the bioactive agent (i.e. a hybridization between the candidate probe and the decoder probe when the bioactive agent is a nucleic acid). After decoding, the DBLs can be removed and the array can be used; however, in some circumstances, for example when the DBL binds to an IBL and not to the bioactive agent, the removal of the DBL is not required (although it may be desirable in some circumstances). In addition, as outlined herein, decoding may be done either before the array is used to in an assay, during the assay, or after the assay.

In one embodiment, a single decoding step is done. In this embodiment, each DBL is labeled with a unique label, such that the number of unique tags is equal to or greater than the number of bioactive

agents (although in some cases, "reuse" of the unique labels can be done, as described herein; similarly, minor variants of candidate probes can share the same decoder, if the variants are encoded in another dimension, i.e. in the bead size or label). For each bioactive agent or IBL, a DBL is made that will specifically bind to it and contains a unique tag, for example one or more fluorochromes. Thus, the identity of each DBL, both its composition (i.e. its sequence when it is a nucleic acid) and its label, is known. Then, by adding the DBLs to the array containing the bioactive agents under conditions which allow the formation of complexes (termed hybridization complexes when the components are nucleic acids) between the DBLs and either the bioactive agents or the IBLs, the location of each DBL can be elucidated. This allows the identification of the location of each bioactive agent; the random array has been decoded. The DBLs can then be removed, if necessary, and the target sample applied.

In a preferred embodiment, the number of unique labels is less than the number of unique bioactive agents, and thus a sequential series of decoding steps are used. To facilitate the discussion, this embodiment is explained for nucleic acids, although other types of bioactive agents and DBLs are useful as well. In this embodiment, decoder probes are divided into **n** sets for decoding. The number of sets corresponds to the number of unique tags. Each decoder probe is labeled in **n** separate reactions with **n** distinct tags. All the decoder probes share the same **n** tags. The decoder probes are pooled so that each pool contains only one of the **n** tag versions of each decoder, and no two decoder probes have the same sequence of tags across all the pools. The number of pools required for this to be true is determined by the number of decoder probes and the **n**. Hybridization of each pool to the array generates a signal at every address. The sequential hybridization of each pool in turn will generate a unique, sequence-specific code for each candidate probe. This identifies the candidate probe at each address in the array. For example, if four tags are used, then 4 X **n** sequential hybridizations can ideally distinguish 4^n sequences, although in some cases more steps may be required. After the hybridization of each pool, the hybrids are denatured and the decoder probes removed, so that the probes are rendered single-stranded for the next hybridization (although it is also possible to hybridize limiting amounts of target so that the available probe is not saturated. Sequential hybridizations can be carried out and analyzed by subtracting pre-existing signal from the previous hybridization).

An example is illustrative. Assuming an array of 16 probe nucleic acids (numbers 1-16), and four unique tags (four different fluors, for example; labels A-D). Decoder probes 1-16 are made that correspond to the probes on the beads. The first step is to label decoder probes 1-4 with tag A, decoder probes 5-8 with tag B, decoder probes 9-12 with tag C, and decoder probes 13-16 with tag D. The probes are mixed and the pool is contacted with the array containing the beads with the attached candidate probes. The location of each tag (and thus each decoder and candidate probe pair) is then determined. The first set of decoder probes are then removed. A second set is added, but this time, decoder probes 1, 5, 9 and 13 are labeled with tag A, decoder probes 2, 6, 10 and 14 are labeled with

tag B, decoder probes 3, 7, 11 and 15 are labeled with tag C, and decoder probes 4, 8, 12 and 16 are labeled with tag D. Thus, those beads that contained tag A in both decoding steps contain candidate probe 1; tag A in the first decoding step and tag B in the second decoding step contain candidate probe 2; tag A in the first decoding step and tag C in the second step contain candidate probe 3; etc. In one embodiment, the decoder probes are labeled in situ; that is, they need not be labeled prior to the decoding reaction. In this embodiment, the incoming decoder probe is shorter than the candidate probe, creating a 5' "overhang" on the decoding probe. The addition of labeled ddNTPs (each labeled with a unique tag) and a polymerase will allow the addition of the tags in a sequence specific manner, thus creating a sequence-specific pattern of signals. Similarly, other modifications can be done, including ligation, etc.

In addition, since the size of the array will be set by the number of unique decoding binding ligands, it is possible to "reuse" a set of unique DBLs to allow for a greater number of test sites. This may be done in several ways; for example, by using some subpopulations that comprise optical signatures. Similarly, the use of a positional coding scheme within an array; different sub-bundles may reuse the set of DBLs. Similarly, one embodiment utilizes bead size as a coding modality, thus allowing the reuse of the set of unique DBLs for each bead size. Alternatively, sequential partial loading of arrays with beads can also allow the reuse of DBLs. Furthermore, "code sharing" can occur as well.

In a preferred embodiment, the DBLs may be reused by having some subpopulations of beads comprise optical signatures. In a preferred embodiment, the optical signature is generally a mixture of reporter dyes, preferably fluoroscent. By varying both the composition of the mixture (i.e. the ratio of one dye to another) and the concentration of the dye (leading to differences in signal intensity), matrices of unique optical signatures may be generated. This may be done by covalently attaching the dyes to the surface of the beads, or alternatively, by entrapping the dye within the bead. The dyes may be chromophores or phosphors but are preferably fluorescent dyes, which due to their strong signals provide a good signal-to-noise ratio for decoding. Suitable dyes for use in the invention include, but are not limited to, fluorescent lanthanide complexes, including those of Europium and Terbium, fluorescein, rhodamine, tetramethylrhodamine, eosin, erythrosin, coumarin, methyl-coumarins, pyrene, Malacite green, stilbene, Lucifer Yellow, Cascade Blue™, Texas Red, and others described in the 6th Edition of the Molecular Probes Handbook by Richard P. Haugland, hereby expressly incorporated by reference.

In a preferred embodiment, the encoding can be accomplished in a ratio of at least two dyes, although more encoding dimensions may be added in the size of the beads, for example. In addition, the labels are distinguishable from one another; thus two different labels may comprise different molecules (i.e. two different fluors) or, alternatively, one label at two different concentrations or intensity.

In a preferred embodiment, the dyes are covalently attached to the surface of the beads. This may be done as is generally outlined for the attachment of the bioactive agents, using functional groups on the surface of the beads. As will be appreciated by those in the art, these attachments are done to minimize the effect on the dye.

In a preferred embodiment, the dyes are non-covalently associated with the beads, generally by entrapping the dyes in the pores of the beads.

Additionally, encoding in the ratios of the two or more dyes, rather than single dye concentrations, is preferred since it provides insensitivity to the intensity of light used to interrogate the reporter dye's signature and detector sensitivity.

In a preferred embodiment, a spatial or positional coding system is done. In this embodiment, there are sub-bundles or subarrays (i.e. portions of the total array) that are utilized. By analogy with the telephone system, each subarray is an "area code", that can have the same tags (i.e. telephone numbers) of other subarrays, that are separated by virtue of the location of the subarray. Thus, for example, the same unique tags can be reused from bundle to bundle. Thus, the use of 50 unique tags in combination with 100 different subarrays can form an array of 5000 different bioactive agents. In this embodiment, it becomes important to be able to identify one bundle from another; in general, this is done either manually or through the use of marker beads, i.e. beads containing unique tags for each subarray.

In alternative embodiments, additional encoding parameters can be added, such as microsphere size. For example, the use of different size beads may also allow the reuse of sets of DBLs; that is, it is possible to use microspheres of different sizes to expand the encoding dimensions of the microspheres. Optical fiber arrays can be fabricated containing features with different fiber diameters or cross-sections; alternatively, two or more fiber optic bundles, each with different cross-sections of the individual fibers, can be added together to form a larger bundle; or, fiber optic bundles with fiber of the same size cross-sections can be used, but just with different sized beads. With different diameters, the largest wells can be filled with the largest microspheres and then moving onto progressively smaller microspheres in the smaller wells until all size wells are then filled. In this manner, the same dye ratio could be used to encode microspheres of different sizes thereby expanding the number of different oligonucleotide sequences or chemical functionalities present in the array. Although outlined for fiber optic substrates, this as well as the other methods outlined herein can be used with other substrates and with other attachment modalities as well.

In addition using the systems described herein, composite arrays are generally described in USSNs 09/256,943, 09/473,904, 09/181,163, 09/606,369, 60/113,968, and PCT/US99/31022, and PCT/US01/20801, all of which are expressly incorporated herein by reference in their entirety.

In a preferred embodiment, the coding and decoding is accomplished by sequential loading of the microspheres into the array. As outlined above for spatial coding, in this embodiment, the optical signatures can be "reused". In this embodiment, the library of microspheres each comprising a different bioactive agent (or the subpopulations each comprise a different bioactive agent), is divided into a plurality of sublibraries; for example, depending on the size of the desired array and the number of unique tags, 10 sublibraries each comprising roughly 10% of the total library may be made, with each sublibrary comprising roughly the same unique tags. Then, the first sublibrary is added to the fiber optic bundle comprising the wells, and the location of each bioactive agent is determined, generally through the use of DBLs. The second sublibrary is then added, and the location of each bioactive agent is again determined. The signal in this case will comprise the signal from the "first" DBL and the "second" DBL; by comparing the two matrices the location of each bead in each sublibrary can be determined. Similarly, adding the third, fourth, etc. sublibraries sequentially will allow the array to be filled.

In a preferred embodiment, codes can be "shared" in several ways. In a first embodiment, a single code (i.e. IBL/DBL pair) can be assigned to two or more agents if the target analytes different sufficiently in their binding strengths. For example, two nucleic acid probes used in an mRNA quantitation assay can share the same code if the ranges of their hybridization signal intensities do not overlap. This can occur, for example, when one of the target sequences is always present at a much higher concentration than the other. Alternatively, the two target sequences might always be present at a similar concentration, but differ in hybridization efficiency.

Alternatively, a single code can be assigned to multiple agents if the agents are functionally equivalent. For example, if a set of oligonucleotide probes are designed with the common purpose of detecting the presence of a particular gene, then the probes are functionally equivalent, even though they may differ in sequence. Similarly, if classes of analytes are desired, all probes for different members of a class such as kinases or G-protein coupled receptors could share a code. Similarly, an array of this type could be used to detect homologs of known genes. In this embodiment, each gene is represented by a heterologous set of probes, hybridizing to different regions of the gene (and therefore differing in sequence). The set of probes share a common code. If a homolog is present, it might hybridize to some but not all of the probes. The level of homology might be indicated by the fraction of probes hybridizing, as well as the average hybridization intensity. Similarly, multiple antibodies to the same protein could all share the same code.

In one embodiment a background subtraction method is used to determine the location of a bead in an array. That is, in contrast to other methods described herein that utilize a bead-based signal to determine the presence of bead in an array, the present invention is directed to a method for determining the location or presence of a bead in an array without the use of a bead-based label. In

one embodiment the invention describes a method to identify the location of beads in the cores of a fiber bundle from by viewing a template image. That is, the method includes generating an image that combines the information in a template image and a foreground image in one single image that does not require hybridization or special chemistry is to identify the location of beads. By template image herein is meant an image where all of features can be identified. By foreground image herein is meant an image where each bead in the array is identified.

In a preferred embodiment, once a population or populations of beads are distributed in an array, an image is created by imaging the bead array with the fluorescent or luminescent material placed at the opposite side of the array relative to the side of the array that contains the detector. That is, for example, when a fiber-optic bundle is used, the fluorescent or luminescent material is shone on oor applied to the beads from the distal end of the array while the image is detected from an optical detector at the proximal end of the fiber-optic bundle.

In one embodiment, when the substrate is a planar substrate and the beads are distributed in wells within the substrate, the array is visualized by detector on one side of the beads. To generate template and foreground images, the substrate is illuminated from the opposite side relative to the detector.

Upon illumination of the bead, light is scattered by the presence of the bead in a well, making the bead location appear relatively dimmer than wells where no beads are present, but still bright enough to identify the well in the array. This method allows for the identification of all cores in the array as well as those wells that contain beads (Figure 13).

Accordingly, the invention provides a method for obtaining an image of an array by illuminating the array, which need not contain a label, and detecting both the array features and the array features that contain beads. As described above, the method includes illuminating the array from the side opposite from the detector and detecting those sites or wells that display reduced illumination relative to other sites. The reduced illumination is an indication of the presence of a bead at the site.

In an alternative embodiment the array is imaged with other sources such as ultraviolet light (i.e. 250-300nm). Again in this embodiment, the array is illuminated from the side opposite to the detector. That is, for example, when the beads are dispersed or distributed in a fiber optic array, the bundle is illuminated from the distal end and detected with a detector located at the proximal end. In one embodiment, the detector is a CCD camera that measures the emission at the proximal end of the array. In an alternative embodiment, multiple CCD cameras are used at the proximal end to measure the emission. In one embodiment, each CCD camera measures intensity at a different wavelength.

In one embodiment, when exciting fluorophores on beads, the fluorophores are excited simultaneously using UV light. Generally, all organic fluorophores contain a benzene-like moiety. Accordingly, they demonstrate absorption in the 250-300 nm region of the spectrum. Once excited, the excitation wavelength is separate and/or distinct from the emission wavelengths. This invention is advantageous over other methods of fluorophore excitation in that previous methods relied on exciting the fluorophore with the l_{max} of the fluorophore. While this frequently resulted in satisfactory discrimination between different types of fluorophores, it prevented spectral imaging of the array because simultaneous excitation could not be achieved without undue overlap between excitation and emission wavelengths. By exciting all the fluorophores simultaneously, simultaneous emission measurement is achieved.

In one aspect this invention makes use of the fact that silica glasses (other than fused silica) display large absorptions in the 250-300 nm region of th spectrum. Accordingly, collection optics can function as efficient light blockers or filters, allowing the use of bright field excitation geometry. Moreover, when the excitation light does not travel far, for example, from the distal end to the proximal end of a fiber optic bundle, it results in reduced or eliminated background light arising from the fluorescence of glass impurities.

In an additional embodiment, the invention involves the simultaneous measurement or detection of a plurality of emission wavelength regimes. That is, for example, when n emission wavelength regimes are involved, these n emission bands are separated by dichroic beam splitters and detected by n-CCD cameras in parallel. This results in speeding up data acquisition by n-fold.

In addition the invention provides a method of abolishing or diminishing residual signal on an array following detection of that signal. That is, generally, following detection of a signal on an array, it is desirable to perform additional analyses on the array. At times, the residual signal from the initial analysis interferes with, and in certain instances can contaminate, subsequent analyses. As such, it is desirable to reduce or diminish any residual signal from a first analysis of an array. In other words, it is desirable to return to the original baseline for detection. As such, the invention provides a method for improving a return to baseline.

In one embodiment the invention includes reducing the residual signal by bleaching the dyes, i.e. fluorophores. In general the method includes performing a first hybridization and detecting the signals. The first signals or labels are then removed. Frequently, residual signals are left on the array; as such the method includes the step of continuing to excite the signals, i.e. dyes or fluorophores, until they bleach or no longer emit a signal. The array can then be used for subsequent analysis.

In a preferred embodiment the first hybridization is a decoding step as described herein. As such, residual signals after decoding hybridization steps are eliminated by bleaching the labels. Methods of

decoding are described in more detail in U.S.S.N 09/748,706 which is expressly incorporated herein by reference. In some embodiments, a stripping step or washing step may be used in combination with the bleaching step. What is important is that the residual signals are substantially reduced prior to subsequent analysis. Preferably at least about 90% or more of the signal is reduced, preferably more than about 99 % of the signal is eliminated, most preferably, greater than about 99.9 % of the signal is eliminated. Following the decoding and removal of residual signals, the signal from the array is returned to baseline. Subsequent analyses to be performed include but are not limited to genotyping, detection of single nucleotide polymorphisms (SNPs), gene expression profiling and the like as are described in more detail in U.S.S.N. 60/244,119, 09/556,463 and 09/553,993 all of which are expressly incorporated herein by reference.

An alternative embodiment of the invention provides a method for improved signal detection by making modifications to standard buffers in which assays are generally interrogated. That is, while a wide range of buffers are commonly used and assays, certain changes can be made to the solution that will positively impact the yield of signal detected from the array. Such modified buffer solutions include solutions that are altered or modified with respect to, but not limited to , alterations in buffer a) temperature, b) viscosity, c) presence of molecular oxygen, or d) volume buffering using hygroscopic solutions. In addition one can adjust the refractive index of the buffer solution to optimize light transfer, for example into the cores of optical fibers of a fiber-optic array.

In a preferred embodiment, the use of hygroscopic salt solutions to control a loss of volume due to evaporation is utilized. Two methods are commonly used to stop evaporation of biological samples. One is simply to cover the sample with oil or enclose it in a tube or seal it in a small space. The second is to keep it in a very high humidity chamber to discourage evaporation. The first method can be costly and cumbersome as the sample is not readily available for handling. Neither of these methods actually buffers the volume of a sample but are merely attempts to discourage evaporation. The present invention relies on a hygroscopic salt (or other hygroscopic material) in the sample to adsorb water from the air or evaporate water to the air. If the concentration of the salt is greater than some equilibrium state with a given humidity of the air the solution will adsorb water and it will release the water if the concentration of the salt is lower than this equilibrium state.

Some salts are less affected by humidity than others. For instance, Sodium Chloride is not as hygroscopic as Sodium Phosphate Dibasic and Potassium salts are one of the most hygroscopic salts available. For this invention, a few suitable hygroscopic salts are for example: barium chloride, calcium chloride, calcium bromide, calcium nitrate, magnesium chloride, magnesium acetate, magnesium nitrate, lithium chloride, ammonium sulfate, ammonium nitrate, ephedrine sulfate, ammonium bromide, ammonium chloride, ammonium iodide, lithium bromide, potassium bisulfate, potassium bromide, potassium chromate, potassium carbonate, potassium iodide, potassium fluoride, potassium nitrate, potassium nitrite, potassium phosphate, potassium thiocyanate, zinc

nitrate, zinc sulfate, cupric sulfate, potassium acetate, potassium citrate, sodium acetate, sodium bisulfate, sodium sulfite, sodium sulfate, sodium dichromate, sodium bromide, sodium phosphate, sodium bromate, sodium fluoride, sodium chlorate, sodium carbonate, sodium iodide, sodium nitrate, sodium nitrite, zinc chloride, phenobarbital sodium, physostigmine sulfate, physostigmine hydrobromide, physostigmine hydrochloride, hydrastine hydrochloride, hydrastine sulfate, thallium sulfate, thallium nitrate, lead nitrate, thallium chloride, sodium chloride, chromium trioxide, phosphoric acid, oxalic acid, and strontium chloride. This list is exemplary and is in no way exhaustive.

With regard to volume, in a preferred embodiment, a saturated solution of LiCl will maintain a constant humidity of 15% in an enclosed space. In the present invention a solution of LiCl would be made and exposed to an environment witch is more humid than 15%. This solution would come to equilibrium with the air and maintain a constant volume after this equilibrium is reached. If the humidity of the air were just slightly greater than 15% the LiCl solution would stabilize at a volume that is just slightly greater than the volume required to make a saturating solution. If a saturated solution of LiCl were exposed to air at 30% humidity the LiCl solution would come to equilibrium at significantly greater volume than the starting volume.

The invention relates to granules of hygroscopic salt in DNA hybridization reactions. The novel media of the invention are able to maintain activity of the biological reaction. It has been found that on formation of a solution of an electrolyte according to the invention, the solution, being hygroscopic, locks in the necessary water molecules to maintain the humidity around the biological reaction.

An example is illustrative. Fluorescence intensity in most molecules is generally thought to increase with decreasing temperature. Without being bound by theory, it is thought that this is a result of decreased frequency of collisions at lower temperatures which reduces the probability for the activation of the excited species by external conversion. Accordingly, when dyed beads, which were distributed on a fiber-optic array, were exposed to decreased temperature, the signal from those beads was increased (Figure 12).

Accordingly, the present invention provides a method for increasing a signal from an array by decreasing the temperature. In a preferred embodiment the array is cooled to a temperature at least below that of room temperature. In a more preferred embodiment the array is cooled to at least 0^0C. The array temperature may be decreased in any manner including but not limited to exposing fiber-immobilized bead arrays to a stream of CO_2 particles. These may have temperatures in the range of -180^0F. Alternatively, the distal end of the fiber bead arrays may be placed into an ice-cooled buffer solution.

Alternatively, the viscosity of the buffer solution may be lowered, thereby increasing this signal. In an alternative embodiment, molecular oxygen may be purged from the buffer solution with, for example, nitrogen gas, resulting in improved signal output.

In one embodiment the invention provides a method for decreasing fluorophore quenching. That is, in some instances, excimer-type fluorophore quenching occurs. Without being bound by theory, it is thought that this quenching is the result of bound fluorophores stacking upon each other on the array surface. Accordingly, the invention includes providing a substance, such as, but not limited cyclodextrins, for reducing the quenching. Cyclodextrins are donut-shaped molecules that are water-soluble, yet possess hydrophobic interiors. Cyclodextrins, especially beta-cyclodextrin, complex various aromatic organic molecules from aqueous solution. The use of cyclodextrins on an array or in a solution applied to an array results in reduced or eliminated fluorophore quenching. In a preferred embodient, cyclodextrin is used to decrease excimer-based quenching of labeled oligos that have hybridized to an array, as described herein, including bead arrays. Thus, the fluorescence intensity of a hybridized bead will increase significantly upon the addition of cyclodextrin to the array solution. As such, the invention also provides a method for increasing signal intensity on an array. In addition, the invention provides a method for increasing assay sensitivity as a result of the increased signal intensity.

In an alternative embodiment, the invention provides a method for background subtraction analysis of an array. In this embodiment, the signal of a bead is determined across a plurality of images. The images are obtained by detecting signals for different emissions, for example, different colors. That is in this embodiment the response of each bead is considered across different images a certain stage. Once acquired, the image with the lowest intensity of the each bead is found and subtracted from the intensity of the bead in the remaining images. As result, the bead response for the lowest channel becomes a zero, and the intensity of the bead in the other channels are reduced accordingly.

The invention finds particular use when background subtraction is automated, for example, when a computer includes code for the automatic determination of the lowest signal of each bead and subtracts that signal from the other images. Accordingly, the invention provides a computer with code programmed for the automatic background subtraction of a bead in an array.

In one embodiment, the array includes at least one internal reference point of known intensity distribution. That is, in addition to microspheres distributed on an array surface, the array includes at least one additional reference feature. In one respect the reference feature or reference point is a

fiducial. The fiducial may be a fiducial bead or fiber as is more fully outlined in U.S.S.N. 09/500,555, and U.S. Application entitled Automated Information Processing in Randomly Ordered Arrays, filed August 9, 2000 (no serial number received), both of which are hereby expressly incorporated by reference.

In a preferred embodiment of the internal reference point is characterized by having a distinct and uniform intensity distribution that is known. Having an internal reference point of known intensity distribution is beneficial in examining an array and allows for correction of fluctuations or irregularities in the array. An example is illustrative.

Images of arrays are often of relatively large-field and high -resolution, and frequently contain artifacts. For example, it is not uncommon for an array image to contain artifacts such as vignetting, which results in reduced signal intensity towards the edge of the array, or non-uniformity as a result of illumination source "hot spots", interference patterns, and the like. By including internal reference points of known intensity in the array, one can measure the signal of the internal reference point and correct the image of the array according to the difference in known signal intensity of the internal reference point and the measured reference point. In a preferred embodiment, when multiple internal reference points are included in the array, the signal intensity of each internal reference point in an image is compared with a pre-determined or known measurement of the distribution of signal intensities for the internal reference point.

In a preferred embodiment, upon comparison of the pre-determined signal intensity of multiple internal reference points with the measured signal from the array, a map of deviation from a reference signal can be plotted. That is the variation of each internal reference point is determined and plotted resulting in a map of variation across the array. Interpolation between points allows for the creation of a matrix comprising a signal non-uniformity correction factor for each location in the array. The matrix finds use in correcting for a variety of defects including image defects including, but not limited to non-uniform illumination of the array, non-uniform photo bleaching and the like.

Accordingly, the invention provides a method for correcting image non-uniformity. The method includes providing an array that includes a substrate with a surface that includes discrete sites such as wells. Beads are distributed in the wells. In addition the array includes at least one internal reference point, but may include a plurality of internal reference points of known signal intensity. The method further includes determining the signal intensity of the internal reference point(s) in the array and comparing the value with the pre-determined signal intensity to obtain the deviation. In an alternative

embodiment the method further includes generating a map of the signal deviation across the array. Once the deviation of the reference points is known, correction of other points on the array is possible.

In an alternative embodiment, each bead in the array contains a reference signal of known intensity. Accordingly, each point or feature of the array can be corrected for alterations in signal intensity. That is, each bead may include a reference label that produces a signal separate from the analytical signal on the bead. This allows a reference signal to be collected for each bead. The reference image may be collected across the entire array before or after obtaining the analytical image. The reference image can then be mapped to the analytical image and used for correction of non-uniformity.

Once made, the compositions of the invention find use in a number of applications. In a preferred embodiment, the compositions are used to probe a sample solution for the presence or absence of a target analyte, including the quantification of the amount of target analyte present. By "target analyte" or "analyte" or grammatical equivalents herein is meant any atom, molecule, ion, molecular ion, compound or particle to be either detected or evaluated for binding partners. As will be appreciated by those in the art, a large number of analytes may be used in the present invention; basically, any target analyte can be used which binds a bioactive agent or for which a binding partner (i.e. drug candidate) is sought.

Suitable analytes include organic and inorganic molecules, including biomolecules. When detection of a target analyte is done, suitable target analytes include, but are not limited to, an environmental pollutant (including pesticides, insecticides, toxins, etc.); a chemical (including solvents, polymers, organic materials, etc.); therapeutic molecules (including therapeutic and abused drugs, antibiotics, etc.); biomolecules (including hormones, cytokines, proteins, nucleic acids, lipids, carbohydrates, cellular membrane antigens and receptors (neural, hormonal, nutrient, and cell surface receptors) or their ligands, etc); whole cells (including procaryotic (such as pathogenic bacteria) and eukaryotic cells, including mammalian tumor cells); viruses (including retroviruses, herpesviruses, adenoviruses, lentiviruses, etc.); and spores; etc. Particularly preferred analytes are nucleic acids and proteins.

In a preferred embodiment, the target analyte is a protein. As will be appreciated by those in the art, there are a large number of possible proteinaceous target analytes that may be detected or evaluated for binding partners using the present invention. Suitable protein target analytes include, but are not limited to, (1) immunoglobulins; (2) enzymes (and other proteins); (3) hormones and cytokines (many of which serve as ligands for cellular receptors); and (4) other proteins.

In a preferred embodiment, the target analyte is a nucleic acid. These assays find use in a wide variety of applications.

In a preferred embodiment, the probes are used in genetic diagnosis. For example, probes can be made using the techniques disclosed herein to detect target sequences such as the gene for nonpolyposis colon cancer, the BRCA1 breast cancer gene, P53, which is a gene associated with a variety of cancers, the Apo E4 gene that indicates a greater risk of Alzheimer's disease, allowing for easy presymptomatic screening of patients, mutations in the cystic fibrosis gene, cytochrome p450s or any of the others well known in the art.

In an additional embodiment, viral and bacterial detection is done using the complexes of the invention. In this embodiment, probes are designed to detect target sequences from a variety of bacteria and viruses. For example, current blood-screening techniques rely on the detection of anti-HIV antibodies. The methods disclosed herein allow for direct screening of clinical samples to detect HIV nucleic acid sequences, particularly highly conserved HIV sequences. In addition, this allows direct monitoring of circulating virus within a patient as an improved method of assessing the efficacy of anti-viral therapies. Similarly, viruses associated with leukemia, HTLV-I and HTLV-II, may be detected in this way. Bacterial infections such as tuberculosis, chlamydia and other sexually transmitted diseases, may also be detected.

In a preferred embodiment, the nucleic acids of the invention find use as probes for toxic bacteria in the screening of water and food samples. For example, samples may be treated to lyse the bacteria to release its nucleic acid, and then probes designed to recognize bacterial strains, including, but not limited to, such pathogenic strains as, *Salmonella, Campylobacter, Vibrio cholerae, Leishmania,* enterotoxic strains of *E. coli*, and Legionnaire's disease bacteria. Similarly, bioremediation strategies may be evaluated using the compositions of the invention.

In a further embodiment, the probes are used for forensic "DNA fingerprinting" to match crime-scene DNA against samples taken from victims and suspects.

In an additional embodiment, the probes in an array are used for sequencing by hybridization.

The present invention also finds use as a methodology for the detection of mutations or mismatches in target nucleic acid sequences. For example, recent focus has been on the analysis of the relationship

between genetic variation and phenotype by making use of polymorphic DNA markers. Previous work utilized short tandem repeats (STRs) as polymorphic positional markers; however, recent focus is on the use of single nucleotide polymorphisms (SNPs), which occur at an average frequency of more than 1 per kilobase in human genomic DNA. Some SNPs, particularly those in and around coding sequences, are likely to be the direct cause of therapeutically relevant phenotypic variants. There are a number of well known polymorphisms that cause clinically important phenotypes; for example, the apoE2/3/4 variants are associated with different relative risk of Alzheimer's and other diseases (see Cordor et al., Science 261(1993). Multiplex PCR amplification of SNP loci with subsequent hybridization to oligonucleotide arrays has been shown to be an accurate and reliable method of simultaneously genotyping at least hundreds of SNPs; see Wang et al., Science, 280:1077 (1998); see also Schafer et al., Nature Biotechnology 16:33-39 (1998). The compositions of the present invention may easily be substituted for the arrays of the prior art.

In a preferred embodiment, the compositions of the invention are used to screen bioactive agents to find an agent that will bind, and preferably modify the function of, a target molecule. As above, a wide variety of different assay formats may be run, as will be appreciated by those in the art. Generally, the target analyte for which a binding partner is desired is labeled; binding of the target analyte by the bioactive agent results in the recruitment of the label to the bead, with subsequent detection.

In a preferred embodiment, the binding of the bioactive agent and the target analyte is specific; that is, the bioactive agent specifically binds to the target analyte. By "specifically bind" herein is meant that the agent binds the analyte, with specificity sufficient to differentiate between the analyte and other components or contaminants of the test sample. However, as will be appreciated by those in the art, it will be possible to detect analytes using binding which is not highly specific; for example, the systems may use different binding ligands, for example an array of different ligands, and detection of any particular analyte is via its "signature" of binding to a panel of binding ligands, similar to the manner in which "electronic noses" work. This finds particular utility in the detection of chemical analytes. The binding should be sufficient to remain bound under the conditions of the assay, including wash steps to remove non-specific binding, although in some embodiments, wash steps are not desired; i.e. for detecting low affinity binding partners. In some embodiments, for example in the detection of certain biomolecules, the dissociation constants of the analyte to the binding ligand will be less than about 10^{-4}-10^{-6} M^{-1}, with less than about 10^{-5} to 10^{-9} M^{-1} being preferred and less than about 10^{-7} -10^{-9} M^{-1} being particularly preferred.

Generally, a sample containing a target analyte (whether for detection of the target analyte or screening for binding partners of the target analyte) is added to the array, under conditions suitable for

binding of the target analyte to at least one of the bioactive agents, i.e. generally physiological conditions. The presence or absence of the target analyte is then detected. As will be appreciated by those in the art, this may be done in a variety of ways, generally through the use of a change in an optical signal. This change can occur via many different mechanisms. A few examples include the binding of a dye-tagged analyte to the bead, the production of a dye species on or near the beads, the destruction of an existing dye species, a change in the optical signature upon analyte interaction with dye on bead, or any other optical interrogatable event.

In a preferred embodiment, the change in optical signal occurs as a result of the binding of a target analyte that is labeled, either directly or indirectly, with a detectable label, preferably an optical label such as a fluorochrome. Thus, for example, when a proteinaceous target analyte is used, it may be either directly labeled with a fluor, or indirectly, for example through the use of a labeled antibody. Similarly, nucleic acids are easily labeled with fluorophor, for example during PCR amplification as is known in the art. Alternatively, upon binding of the target sequences, a hybridization indicator may be used as the label. Hybridization indicators preferentially associate with double stranded nucleic acid, usually reversibly. Hybridization indicators include intercalators and minor and/or major groove binding moieties. In a preferred embodiment, intercalators may be used; since intercalation generally only occurs in the presence of double stranded nucleic acid, only in the presence of target hybridization will the label light up. Thus, upon binding of the target analyte to a bioactive agent, there is a new optical signal generated at that site, which then may be detected.

Alternatively, in some cases, as discussed above, the target analyte such as an enzyme generates a species that is either directly or indirectly optical detectable.

Furthermore, in some embodiments, a change in the optical signature may be the basis of the optical signal. For example, the interaction of some chemical target analytes with some fluorescent dyes on the beads may alter the optical signature, thus generating a different optical signal.

As will be appreciated by those in the art, in some embodiments, the presence or absence of the target analyte may be done using changes in other optical or non-optical signals, including, but not limited to, surface enhanced Raman spectroscopy, surface plasmon resonance, radioactivity, etc.

In one embodiment the invention provides a method for improving the efficiency of signal generation from an array. In some embodiments a signal is generated from a localized, well-defined region of an array such as a discrete site or well as described herein. Certain useful assays are based on the

generation of a diffusible signal in liquid phase. Examples include, but are not limited to, pyrosequencing and the ReadIT assay (Promega), both based on luminescence generated by luciferase. In order to carry out these and other reactions in parallel on an array format, the present invention provides ways to prevent loss or mixing of signals by diffusion in the time-frame required for detection. In one embodiment this is accomplished, for example, by physically separating each assay. That is, assays are performed in a well within which is at least one and in some instances more, discrete site or well as described herein.

An alternative approach is to read the signal in real-time, before it is able to diffuse away (e.g. Michael, K.L., Taylor, L.C., Schultz, S.L. & Walt, D.R. Randomly ordered addressable high-density optical sensor arrays. *Analytical Chemistry* 70:1242-1248 (1998), incorporated herein by reference. This embodiment makes use of a patterned substrate that includes sites or wells into which beads are distributed and also contains a pattern that surrounds the beads, as described herein (Fig. I5). As previously noted, the region surrounding each site includes functionalized sites for the attachment of detection moieties. Examples of detection moieties include fluorescence resonance energy transfer (FRET) molecules that may be directly or indirectly attached to the functionalized site, reporter molecules including dyes and enzymes. In a preferred embodiment the detection moiety includes reporter enzymes such as but not limited to firefly luciferase, and b-galactosidase, In a preferred embodiment the region around each discrete site or bead is used to immobilize one or more detection enzymes. Luciferase and/or sulfurylase are particularly preferred. They find particular use in pyrosequencing.

Accordingly, the method includes providing a substrate with at least a first assay region comprising at least a first site comprising a bioactive agent around which are functionalized for attachment of a detection moiety. In one embodiment the array comprises the detection moiety attached to the substrate. In a preferred embodiment the site comprising a bioactive agent is a microsphere comprising a bioactive agent, wherein the microsphere is distributed in a well on the array. The method further includes contacting the array with a sample containing a target analyte and monitoring the production of a product released from the microsphere by detecting a signal produced by the detection moieties. A signal produced by the detection moieties is an indication of an interaction between the target analyte and the bioactive agent on the microsphere.

The method finds particular use in analyzing enzymatic assays as enzymes release a diffusable product(s) that is detected by detection moieties including various reporter molecules as described herein.

Again, as outlined above for decoding, the assay for the presence or absence of a target analyte utilizes sequential processing of data images using a computer system. Thus, in a preferred embodiment, a first data image of a random array is acquired using an acquisition module of the computer system. This initial data image may be decoded, i.e. the location of some or all of the bioactive agents may be known, or decoding may occur either during or after the assay. A registration module of the computer system is used to create a registered first data image, using either an exogeneous fiducial or a fiducial template generated by acquiring a template data image as outlined herein, for example by evening illuminating the array. The sample is then added to the array, and a second data image is acquired using the acquisition module. The fiducial and registration module are then used to create a registered second data image. A comparison module of the computer system is then used to compare the registered data images to determine the presence or absence of said target analyte.

The assays may be run under a variety of experimental conditions, as will be appreciated by those in the art. A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding. Various blocking and washing steps may be utilized as is known in the art.

In a preferred embodiment, two-color competitive hybridization assays are run. These assays can be based on traditional sandwich assays. The beads contain a capture sequence located on one side (upstream or downstream) of the SNP, to capture the target sequence. Two SNP allele-specific probes, each labeled with a different fluorophor, are hybridized to the target sequence. The genotype can be obtained from a ratio of the two signals, with the correct sequence generally exhibiting better binding. This has an advantage in that the target sequence itself need not be labeled. In addition, since the probes are competing, this means that the conditions for binding need not be optimized. Under conditions where a mismatched probe would be stably bound, a matched probe can still displace it. Therefore the competitive assay can provide better discrimination under those conditions. Because many assays are carried out in parallel, conditions cannot be optimzed for every probe simultaneously. Therefore, a competitive assay system can be used to help compensate for non-optimal conditions for mismatch discrimination.

In a preferred embodiment, dideoxynucleotide chain-termination sequencing is done using the compositions of the invention. In this embodiment, a DNA polymerase is used to extend a primer using fluorescently labeled ddNTPs or other chain terminating nucleotides. The 3' end of the primer is located adjacent to the SNP site. In this way, the single base extension is complementary to the sequence at the SNP site. By using four different fluorophors, one for each base, the sequence of the SNP can be deduced by comparing the four base-specific signals. This may be done in several ways. In a first embodiment, the capture probe can be extended; in this approach, the probe must either be synthesized 5'-3' on the bead, or attached at the 5' end, to provide a free 3' end for polymerase extension. Alternatively, a sandwich type assay can be used; in this embodiment, the target is captured on the bead by a probe, then a primer is annealed and extended. Again, in the latter case, the target sequence need not be labeled. In addition, since sandwich assays require two specific interactions, this provides increased stringency which is particularly helpful for the analysis of complex samples.

In addition, primer extension is possible; extension of a primer bound to template in liquid phase is followed by capture of the extended primer on the array.

In addition, when the target analyte and the DBL both bind to the agent, it is also possible to do detection of non-labelled target analytes via competition of decoding.

In a preferred embodiment, the methods of the invention are useful in array quality control. Prior to this invention, no methods have been described that provide a positive test of the performance of every probe on every array. Decoding of the array not only provides this test, it also does so by making use of the data generated during the decoding process itself. Therefore, no additional experimental work is required. The invention requires only a set of data analysis algorithms that can be encoded in software.

The quality control procedure can identify a wide variety of systematic and random problems in an array. For example, random specks of dust or other contaminants might cause some sensors to give an incorrect signal-this can be detected during decoding. The omission of one or more agents from multiple arrays can also be detected. An advantage of this quality control procedure is that it can be implemented immediately prior to the assay itself, and is a true functional test of each individual sensor. Therefore any problems that might occur between array assembly and actual use can be detected. In applications where a very high level of confidence is required, and/or there is a significant

chance of sensor failure during the experimental procedure, decoding and quality control can be conducted both before and after the actual sample analysis.

In a preferred embodiment, the arrays can be used to do reagent quality control. In many instances, biological macromolecules are used as reagents and must be quality controlled. For example, large sets of oligonucleotide probes may be provided as reagents. It is typically difficult to perform quality control on large numbers of different biological macromolecules. The approach described here can be used to do this by treating the reagents (formulated as the DBLs) as variable instead of the arrays.

In a preferred embodiment, the methods outlined herein are used in array calibration. For many applications, such as mRNA quantitation, it is desirable to have a signal that is a linear response to the concentration of the target analyte, or, alternatively, if non-linear, to determine a relationship between concentration and signal, so that the concentration of the target analyte can be estimated. Accordingly, the present invention provides methods of creating calibration curves in parallel for multiple beads in an array. The calibration curves can be created under conditions that simulate the complexity of the sample to be analyzed. Each curve can be constructed independently of the others (e.g. for a different range of concentrations), but at the same time as all the other curves for the array. Thus, in this embodiment, the sequential decoding scheme is implemented with different concentrations being used as the code "labels", rather than different fluorophores. In this way, signal as a response to concentration can be measured for each bead. This calibration can be carried out just prior to array use, so that every probe on every array is individually calibrated as needed.

In a preferred embodiment, the methods of the invention can be used in assay development as well. Thus, for example, the methods allow the identification of good and bad probes; as is understood by those in the art, some probes do not function well because they do not hybridize well, or because they cross-hybridize with more than one sequence. These problems are easily detected during decoding. The ability to rapidly assess probe performance has the potential to greatly reduce the time and expense of assay development.

Similarly, in a preferred embodiment, the methods of the invention are useful in quantitation in assay development. Major challenges of many assays is the ability to detect differences in analyte concentrations between samples, the ability to quantitate these differences, and to measure absolute concentrations of analytes, all in the presence of a complex mixture of related analytes. An example of this problem is the quantitation of a specific mRNA in the presence of total cellular mRNA. One approach that has been developed as a basis of mRNA quantitation makes use of a multiple match

and mismatch probe pairs (Lockhart et al., 1996), hereby incorporated by reference in its entirety. While this approach is simple, it requires relatively large numbers of probes. In this approach, a quantitative response to concentration is obtained by averaging the signals from a set of different probes to the gene or sequence of interest. This is necessary because only some probes respond quantitatively, and it is not possible to predict these probes with certainty. In the absence of prior knowledge, only the average response of an appropriately chosen collection of probes is quantitative. However, in the present invention, that can be applied generally to nucleic acid based assays as well as other assays. In essence, the approach is to identify the probes that respond quantitatively in a particular assay, rather than average them with other probes. This is done using the array calibration scheme outlined above, in which concentration-based codes are used. Advantages of this approach include: fewer probes are needed; the accuracy of the measurement is less dependent on the number of probes used; and that the response of the sensors is known with a high level of certainty, since each and every sequence can be tested in an efficient manner. It is important to note that probes that perform well are selected empirically, which avoids the difficulties and uncertainties of predicting probe performance, particularly in complex sequence mixtures. In contrast, in experiments described to date with ordered arrays, relatively small numbers of sequences are checked by performing quantitative spiking experiments, in which a known mRNA is added to a mixture.

Accordingly, the present invention also includes an array as described above and a detector. In a preferred embodiment the array is in the detector. In a particularly preferred embodiment the substrate is a microscope slide with fiber optic bundles forming the assay locations and array locations and this array is in the detector.

All references cited herein are incorporated by reference in their entirety.

Examples

Example 1 Improved Signal Detection from arrays

The ability for any analytical system, optical or otherwise, to detect a change in signal is dictated by the background and noise associated with that signal. Exploring different ways to improve the signal to background ratio by either amplifying the signal, reducing the background, or both, is thus a critical area of research during the development of any type of analytical detection system.

One of the primary sources of background in optical microarray systems is the intrinsic fluorescence of the array substrate. Generally, the fluorescence of microspheres immobilized at the distal tip of the imaging fiber bundle is imaged from the proximal end of the bundle. While this approach has a number of advantages, most importantly being the remote sensing capability and the ease of sample interface, the background of the measurement will necessarily include any fluorescence originating from the fiber core glass itself. Since each fiber element is its own waveguide, it is particularly susceptible to generating high fluorescence readings on the detector since the fluorescence of the glass constituents as well as any contaminants present at the core-clad interface will be captured and propagated down the fiber and measured by the detector. Conversely, if one turns the fiber around and images the bead array directly, the background is found to be slightly reduced (most likely due to the fact that the focal plane is no longer placed on the glass itself, but rather on the beads in the wells, and thus collection of core fluorescence is not as efficient). This effect is shown in Figure 1.

In this configuration, when it is not necessary to view through the fiber, a non-fluorescent coating such as a thin metal film can be placed over the etched array that blocks the excitation light from hitting the fluorescent substrate underneath, thereby effectively reducing the background of the array. This is demonstrated in Figure 2.

For this experiment, fluorescein-labeled silica beads were loaded into two different etched fiber bundles: one coated with a thin palladium film (via vapor-deposition), the other uncoated. The average intensities of a subset of beads and empty cores were measured for each fiber and graphed in Figure 2a. The results indicate a substantial reduction in background of the metal-coated fiber as compared to the uncoated fiber, resulting in a 10-fold improvement of the signal-to-background ratio.

This technique could be used to improve sensitivity of a wide range of assays, including SNP genotyping, small molecule screening, immunoassays, peptide or protein assays, enzymatic assays, and any other chemical or biological assay that can be performed in an optical microarray format.

Example 2 Improved Signal Output from Beads Distributed in Wells Compared to Beads on a Flat Surface

To examine the effect on signal output of beads distributed in wells as compared to beads placed on a flat surface, dyed beads were distributed on a chip that contained both wells and a flat surface. Upon detection of the signal of beads, it was evident that those in the wells produced a more intense signal as compared to beads on a flat surface. Without being bound by theory, it is thought that in placing beads in a recessed well of any kind may bring an added advantage over a flat surface; namely the capture of additional photons through a "micro-reflector" effect. The results of this are depicted in Figure 8.

Example 3

To examine the effect of buffer solutions on signal output, experiments were performed that involved lowering the temperature of dyed beads (both silica and glycidyl methacrylate (GMA) beads) by two methods: 1) exposing the fiber-immobilized bead arrays to a stream of CO_2 particles with temperatures of approximately -180F, and 2) placing the distal end of the fiber bead arrays into an ice-cooled buffer solution. The results are displayed in Figure 12.

In those experiments, the CO_2-cooled fiber arrays showed approximately a two-fold enhancement in fluorescence over room temperature arrays.

We claim:

1. A composition comprising:

a. a substrate with a surface comprising wells, wherein said wells comprise concave floors; and

b. a population of microspheres randomly distributed in said wells.

2. A composition according to claim 1, wherein said substrate is a fiber optic bundle.

3. A composition according to claim 1, wherein said substrate is selected from the group consisting of glass and plastic.

4. A composition according to claim 1, 2 or 3, wherein said microspheres further comprise a bioactive agent.

5. A composition according to claim 1, 2, 3, or 4 wherein said population of microspheres comprises at least first and second subpopulations.

6. A method of making an array comprising:

a. providing a surface comprising wells, wherein said wells comprise concave floors; and

b. randomly distributing a population of microspheres in said wells.

7. The method according to claim 6 wherein the surface area of said microspheres that is in contact with said wells is increased when compared with convex shaped wells.

8. A composition according to claim 7, wherein said substrate is a fiber optic bundle.

9. A composition according to claim 7, wherein said substrate is selected from the group consisting of glass and plastic.

10. A composition according to claim 7, 8, or 9, wherein said microspheres further comprise a bioactive agent.

11. A composition according to claim 7, 8, 9, or 10, wherein said population of microspheres comprises at least first and second subpopulations.

12. A method of imaging an array comprising:

a. providing an array comprising a plurality of discrete sites;

b. applying said array to a detector;

c. imaging said array with said detector, wherein at least a first and a second discrete site are imaged within a first and a second pixel, respectively.

13.. A method according to claim 12, wherein each of said discrete sites is imaged within a single discrete pixel.

14. A method according to claim 12 and 13, wherein said array is an ordered array.

15. A method according to claim 12 and 13, wherein said array is a random array.

16. A method according to claim 15, wherein said array comprises a population of microspheres randomly distributed in said discrete sites.

17. A method according to claim 16, wherein said population of microspheres comprise a first and a second subpopulation.

18. A method according to claim 16, wherein said population of microspheres comprise bioactive agents.

19. A method for maintaining humidity during of an array comprising:

a. providing an array comprising:

i) a substrate with surface comprising discrete sites; and

ii) a population of microspheres randomly distributed in said sites, wherein each of said microspheres comprises a bioactive agent;

b. contacting said array with a solution comprising a hygroscopic salt in an initial concentration to form a hygroscopic solution, whereby said hygroscopic solution maintains said humidity.

20. A method according to claim 19 wherein said population of microspheres comprise a first and a second subpopulation.

21. A method according to claim 19 wherein said hygroscopic salt is selected from the group consisting of potassium bisulfate, potassium bromide, potassium chromate, potassium carbonate, and potassium iodide.

22. A method according to claim 19 wherein said salt is LiCl.

23. A method according to claim 19 wherein said substrate is a fiber optic bundle.

24. A method according to claim 19 wherein said substrate is selected from glass and plastic.

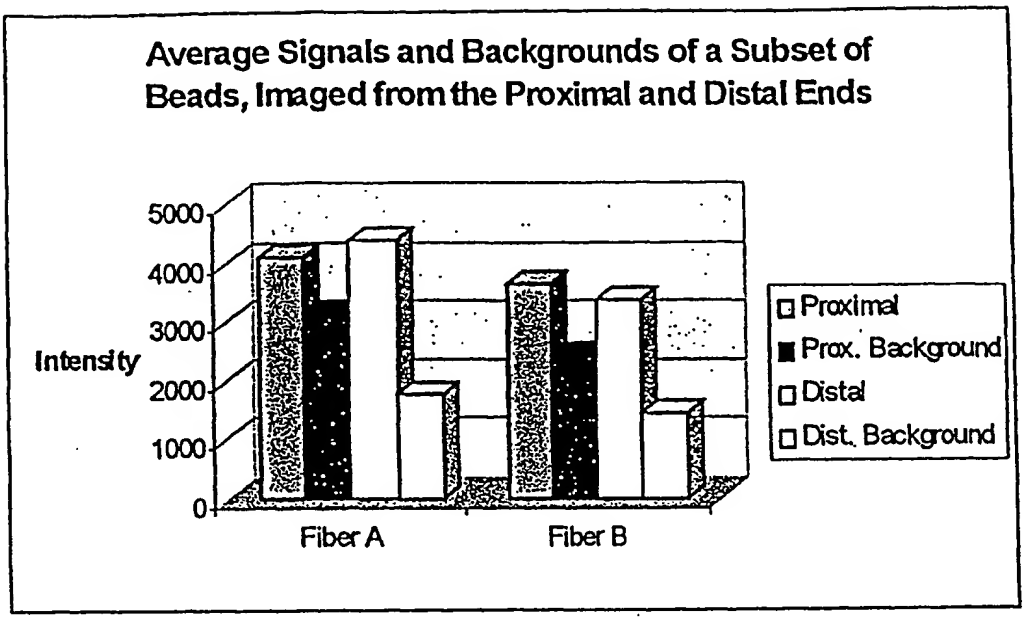

Average Signals and Backgrounds of a Subset of Beads, Imaged from the Proximal and Distal Ends
Intensity
5000
4000
3000
2000
1000
0
Fiber A
Fiber B
Proximal
Prox. Background
Distal
Dist. Background

FIGURE 1

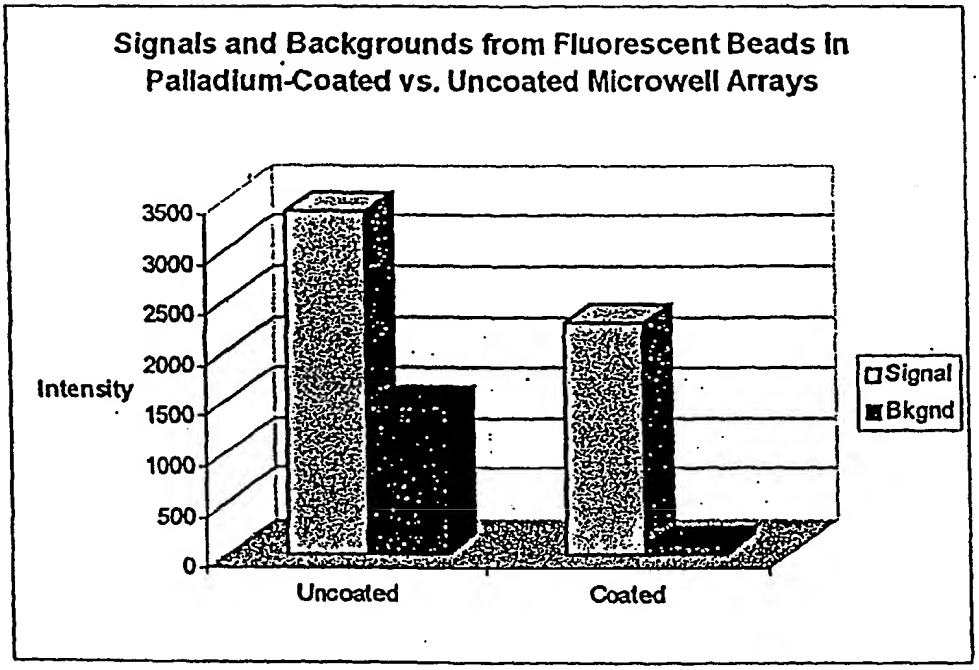

Signals and Backgrounds from Fluorescent Beads in Palladium-Coated vs. Uncoated Microwell Arrays
Intensity
3500
3000
2500
2000
1500
1000
500
0
Uncoated
Coated
Signal
Bkgnd

FIGURE 2a

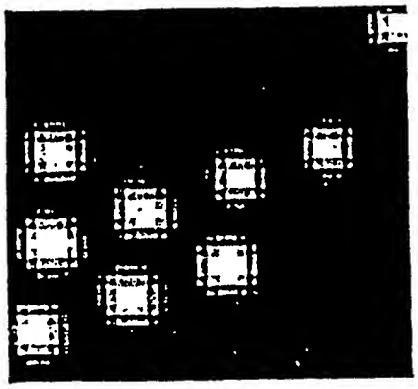


FIGURE 2b

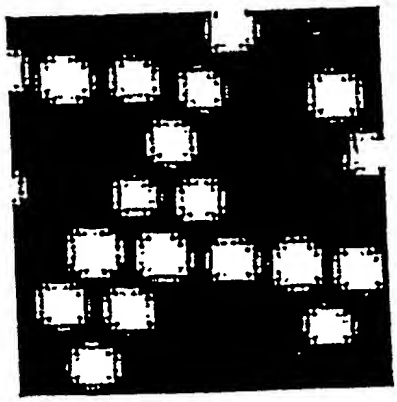


FIGURE 2c

FIGURE 3

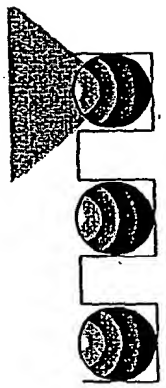


FIGURE 4a

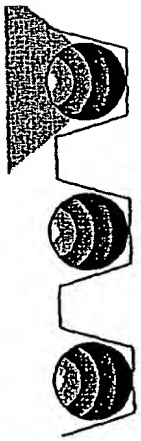


FIGURE 4b

FIGURE 4c

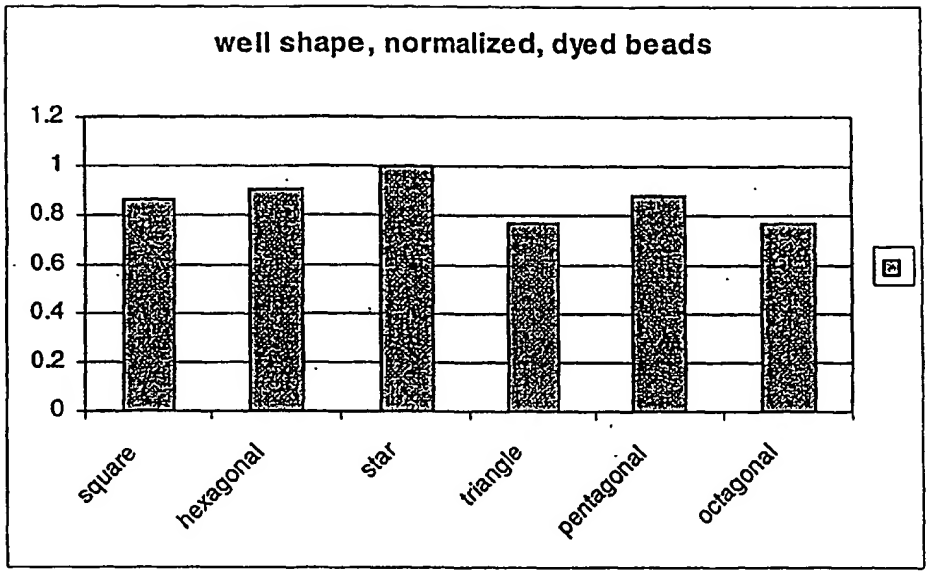

well shape, normalized, dyed beads
1.2
1
0.8
0.6
0.4
0.2
0
square
hexagonal
star
triangle
pentagonal
octagonal

FIGURE 5

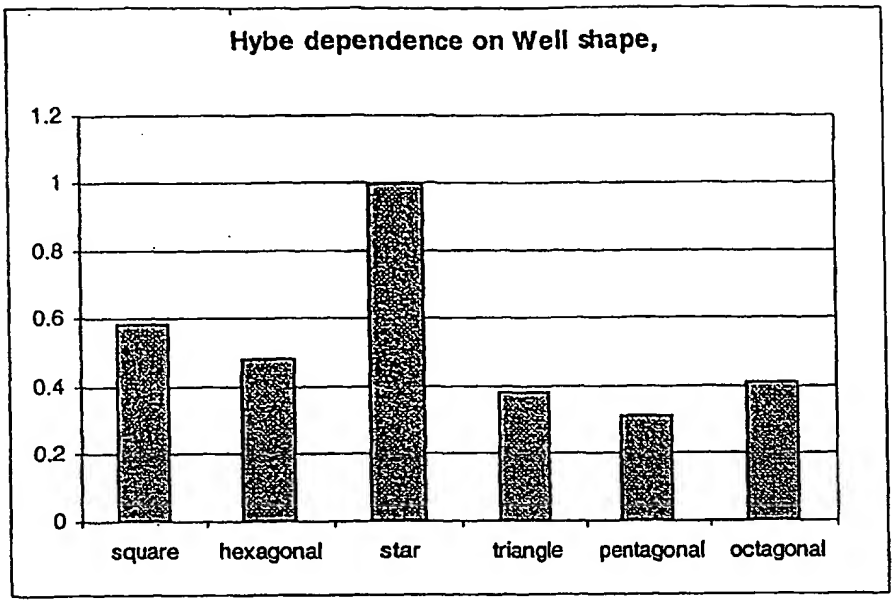


FIGURE 6

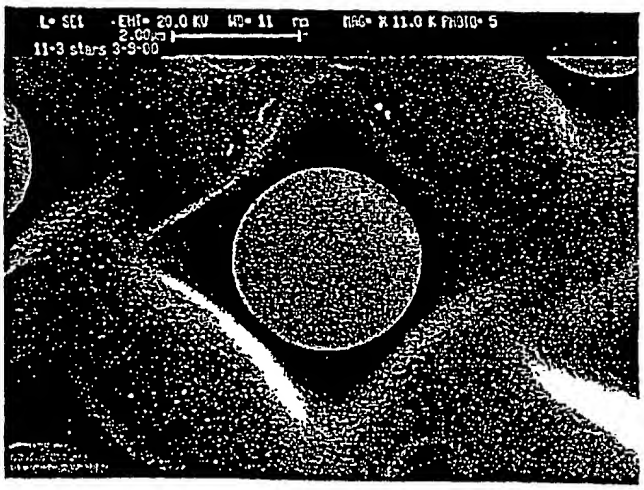


FIGURE 7

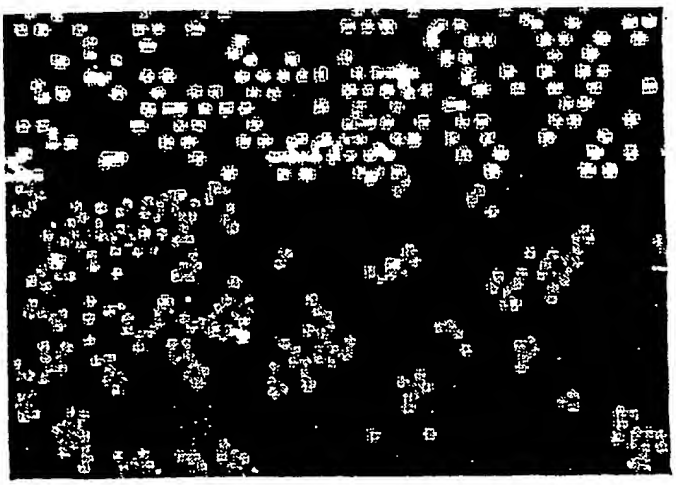


FIGURE 8a

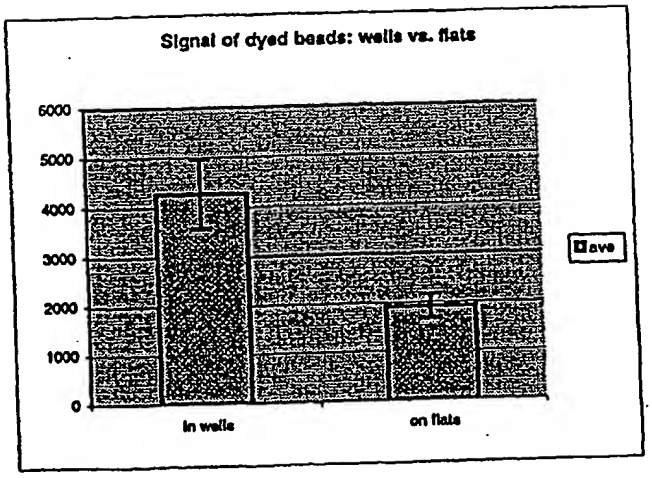


FIGURE 8b

Concept 1.
Unlabeled Analyte
T
C
A
A
T
Hybridization
F
A
G
T
A
T
Bead
Fluorescence State One
A
G
T
F
A
T
C
T
A
A
T
Bead
Fluorescence State Two

FIGURE 9

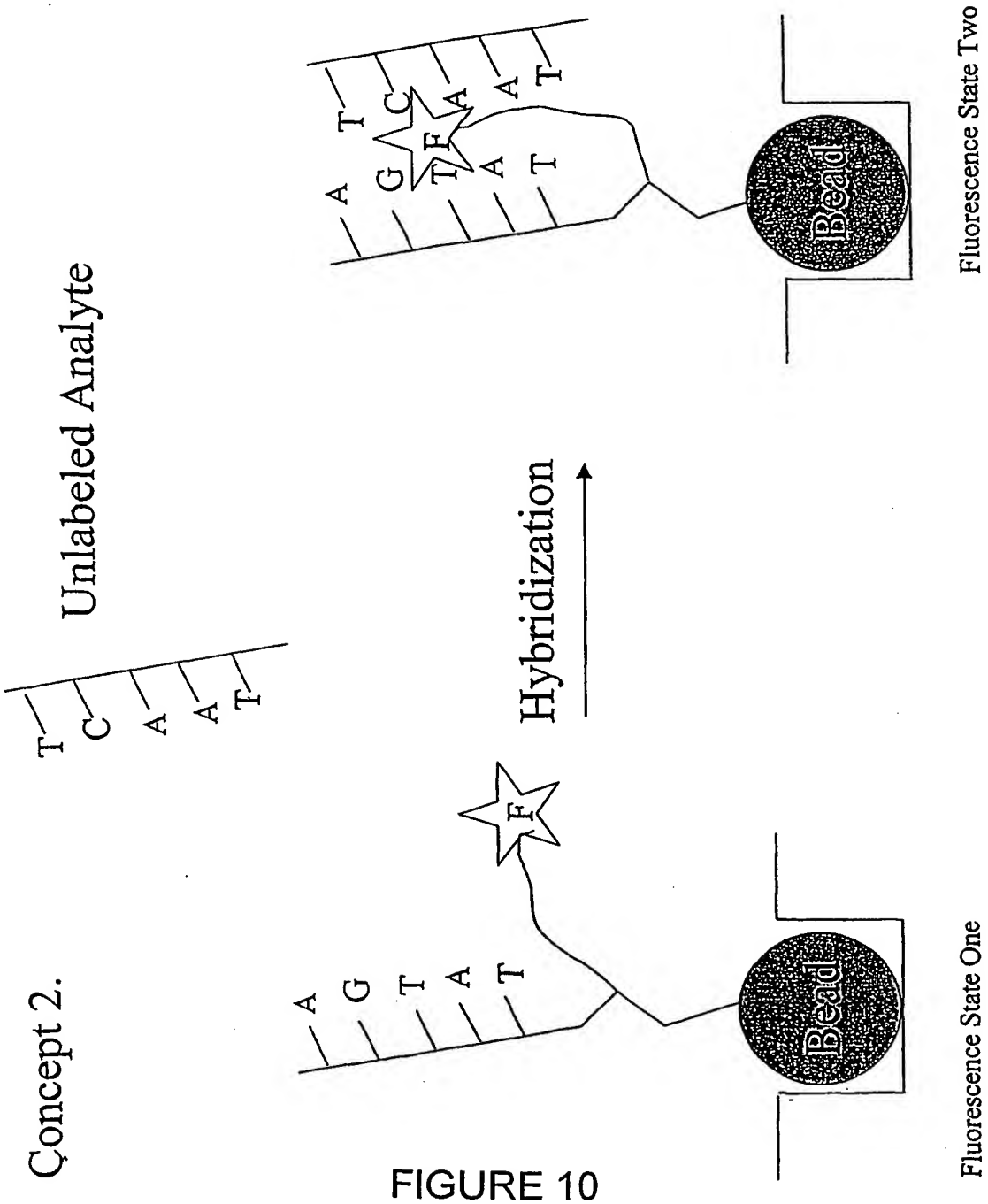

Concept 2.
Unlabeled Analyte
T
C
A
A
T
Hybridization
A
G
T
A
T
F
Bead
Fluorescence State One
A
G
T
F
A
A
T
Bead
Fluorescence State Two

FIGURE 10

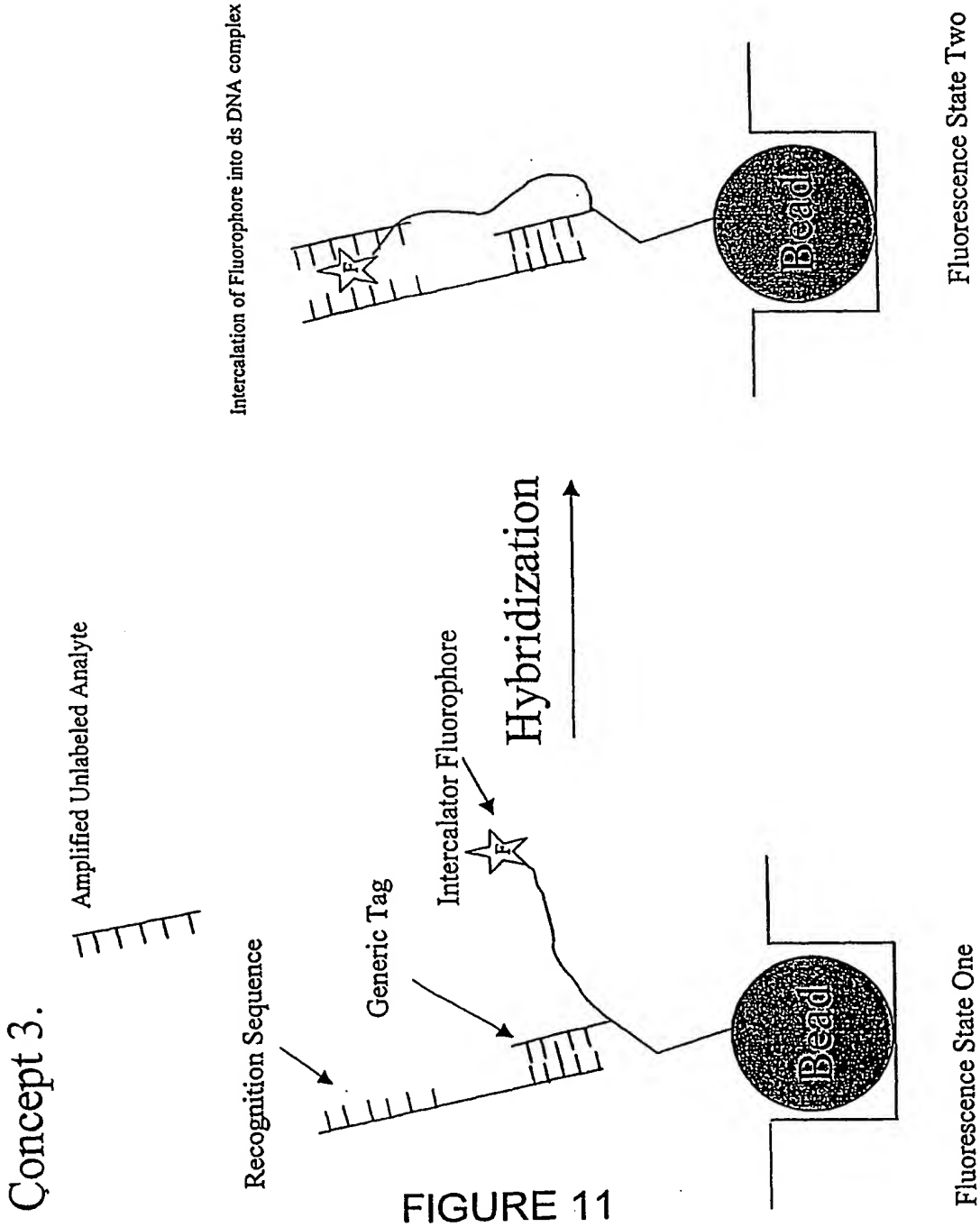

Concept 3.
Amplified Unlabeled Analyte
Recognition Sequence
Generic Tag
Intercalator Fluorophore
F
Hybridization
Intercalation of Fluorophore into ds DNA complex
F
Bead
Bead
Fluorescence State One
Fluorescence State Two

FIGURE 11

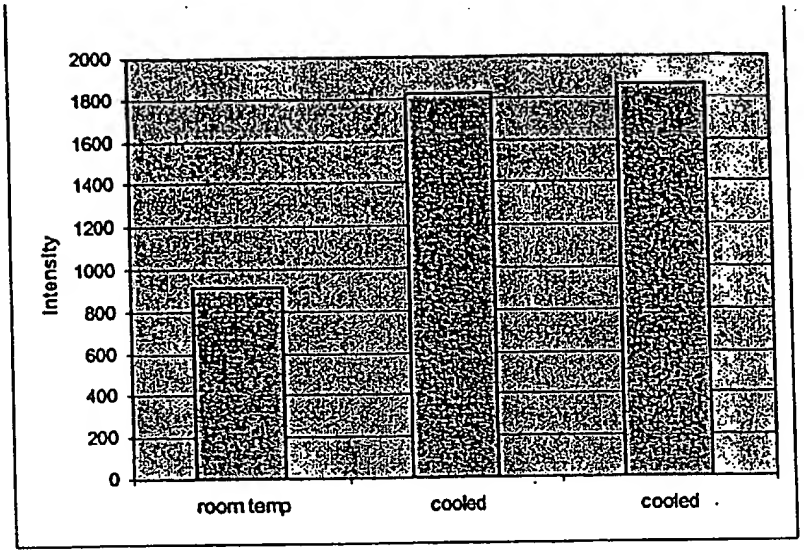


FIGURE 12a

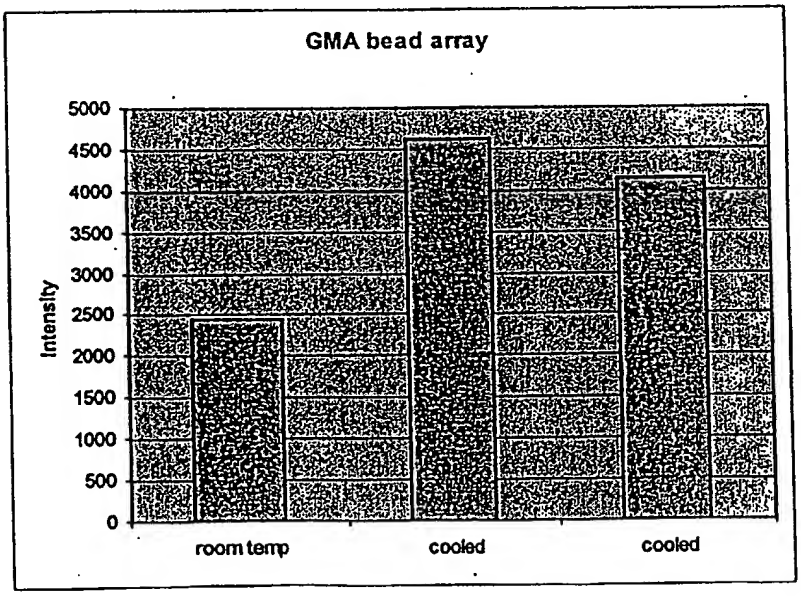


FIGURE 12b

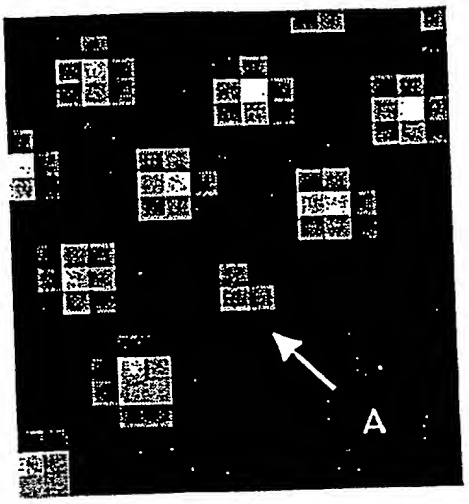


FIGURE 13a

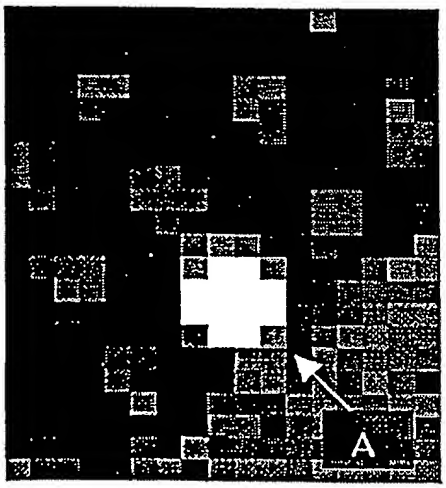


FIGURE 13b

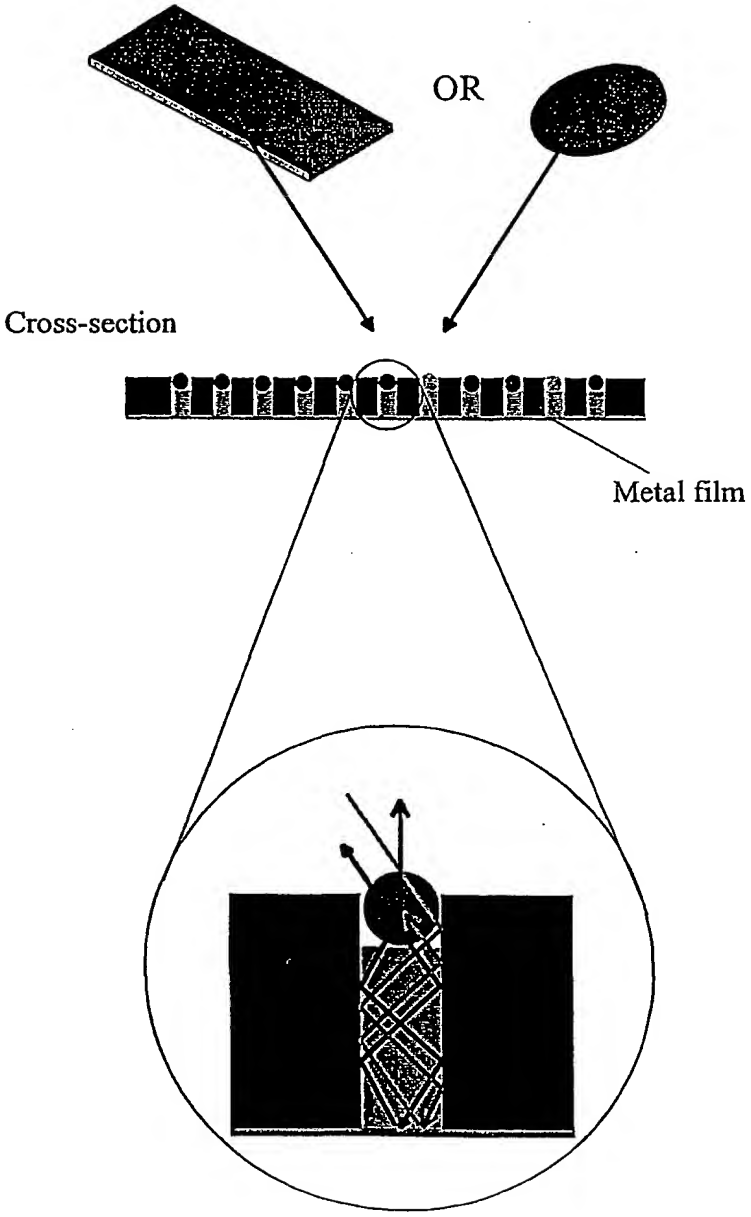


FIGURE 14

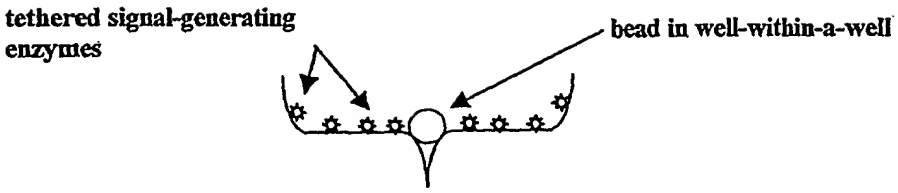

tethered signal-generating enzymes
bead in well-within-a-well

Figure 15

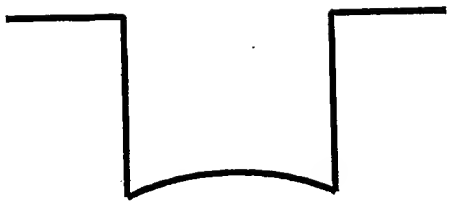


Figure 16A

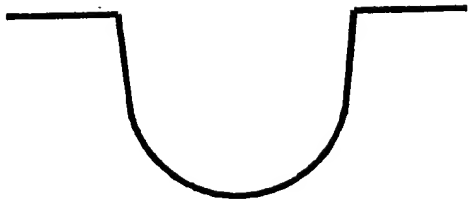


Figure 16B

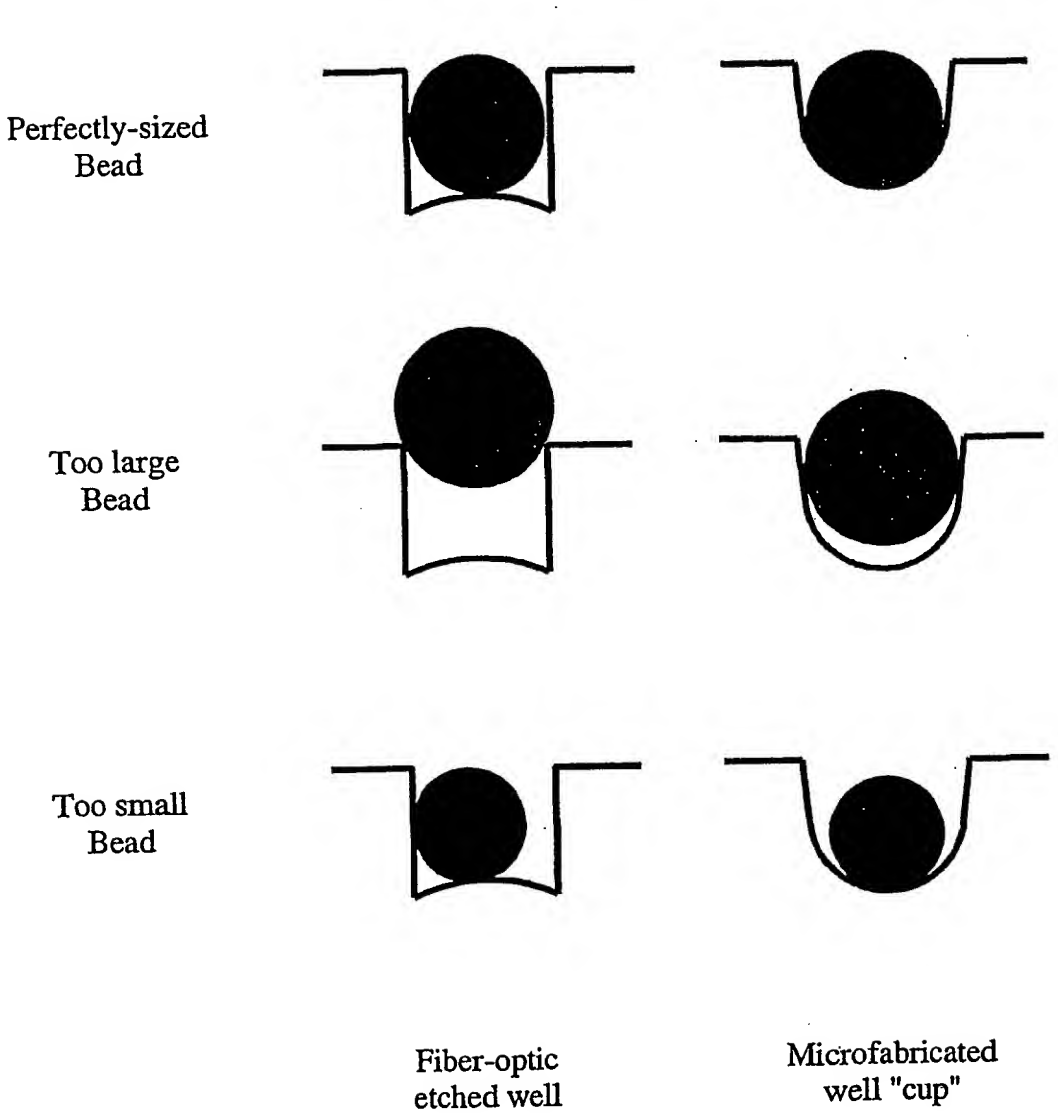

Perfectly-sized Bead
Too large Bead
Too small Bead
Fiber-optic etched well
Microfabricated well "cup"

FIGURE 17

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2002 (12.12.2002)

PCT

(10) International Publication Number
WO 02/099982 A3

(51) International Patent Classification[7]: C12M 1/36, C12N 11/16, G01N 15/06, C12Q 1/68, G02B 6/26

(21) International Application Number: PCT/US02/06327

(22) International Filing Date: 1 March 2002 (01.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/272,803 1 March 2001 (01.03.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/272,803 (CIP)
Filed on 1 March 2001 (01.03.2001)

(71) Applicant *(for all designated States except US)*: **ILLUMINA, INC.** [US/US]; 9885 Towne Centre Drive, San Diego, CA 92121-1975 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DICKINSON, Todd** [US/US]; 3435 Lebon Drive #1133, San Diego, CA 92122 (US). **OLIPHANT, Arnold** [US/US]; 15518 Markar Road, Poway, CA 92064-2315 (US).

(74) Agents: **BREZNER, David, J.** et al.; Flehr, Hohbach, Test, Albritton & Herbert LLP, Suite 3400, Four Embarcadero Center, San Francisco, CA 94111-4187 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
20 November 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/099982 A3

(54) Title: METHODS FOR IMPROVING SIGNAL DETECTION FROM AN ARRAY

(57) Abstract: The invention relates methods of improving signal detection from an array and methods for background subtraction in an array. The invention provides for novel array compositions including arrays with wells with different shapes, or surfaces coated with reflective or selectively absorptive coatings. In addition, the array include a signal transducer element.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/06327

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12M 1/36; C12N 11/16; G01N 15/06; C12Q 1/68; G02B 6/26
US CL : 435/4, 174, 293.1, 287.2; 422/68.1; 385/12
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/4, 174, 293.1, 287.2; 422/68.1; 385/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,023,540 A (WALT et al) 08 February 2000 (08.02.2000) columns 3-4 and 11-14 and Fig. 5.	1-4, 6-10, 12-18
Y		19-24
Y	US 5,412,087 A (MCGALL et al) 02 May 1995 (02.05.1995) columns 6 and 16.	1-24
Y	US 6,045,996 A (CRONIN et al) 04 April 2000 (04.04.2000) columns 3-6, 8 and 10.	1-24

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 November 2002 (18.11.2002)	22 MAY 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer BJ Forman Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/06327

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claim Nos.: 5 and 11
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 December 2002 (19.12.2002)

PCT

(10) International Publication Number
WO 02/101358 A2

(51) International Patent Classification[7]: G01N

(21) International Application Number: PCT/US02/18634

(22) International Filing Date: 11 June 2002 (11.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/297,609	11 June 2001 (11.06.2001)	US
09/915,231	24 July 2001 (24.07.2001)	US
60/311,271	9 August 2001 (09.08.2001)	US

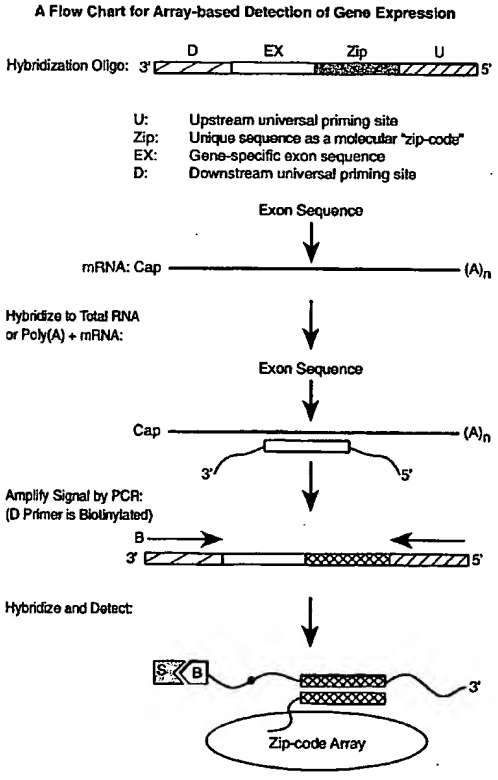
(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/915,231 (CIP)
Filed on 24 July 2001 (24.07.2001)

(71) Applicant *(for all designated States except US)*: **ILLUMINA, INC.** [US/US]; 9885 Towne Centre Drive, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **FAN, Jian-Bing** [CN/US]; 5247 Pearlman Way, San Diego, CA 92130 (US). **STUELPNAGEL, John, R.** [US/US]; 38 Briggs Avenue, Encinitas, CA 92024 (US). **CHEE, Mark, S.** [AU/US]; 155 15th Street, No. 22, Del Mar, CA 92014 (US).

[Continued on next page]

(54) Title: MULTIPLEXED DETECTION METHODS



(57) Abstract: The present invention permits highly multiplexed detection of target molecules in a sample. The method includes contacting target analytes with a composition comprising an amplification enzyme and first and second target probes. The first and second target probes comprise first and second bioactive agent respectively, that specifically bind to the first and second target molecules. The probes may also comprise first and second adapter sequences, respectively, such and that the first adapter sequence identifies the first target molecule and the second adapter sequence identifies the second target molecule, and first and section upstream universal priming sequences, respectively. Finally, following amplification the method includes detecting the first and second amplicons, whereby the first and second target molecules, respectively are detected.

WO 02/101358 A2

(74) **Agents: BREZNER, David, J.** et al.; Dorsey & Whitney LLP, 4 Embarcadero Center, Suite 3400, San Francisco, CA 94111 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

(48) **Date of publication of this corrected version:**
27 March 2003

(15) **Information about Correction:**
see PCT Gazette No. 13/2003 of 27 March 2003, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MULTIPLEXED DETECTION METHODS

The present application claims the benefit of application U.S.S.N.s 60/297,609, filed June 11, 2001, and 60/311,271, filed August 9, 2001, both of which are expressly incorporated herein by reference, and is a continuing application of 09/915,231, filed July 24, 2001, which is expressly incorporated herein by reference.

FIELD OF THE INVENTION

The present invention is directed to sensitive and accurate multiplexed assays for target analyte detection.

BACKGROUND OF THE INVENTION

The detection of various target analytes or molecules is an important tool for a variety of application including diagnostic medicine, molecular biology research and detection of contaminants, to name a few. While method of detecting different analytes has evolved, the ability to detect numerous target analytes simultaneously has proven difficult. Detection of multiple proteins, for example has been limited to conventional electrophoresis assays or immunoassays. There has not been a significant multiplexed protein detection assay or method.

The detection of specific nucleic acids is an important tool for diagnostic medicine and molecular biology research. Gene probe assays currently play roles in identifying infectious organisms such as bacteria and viruses, in probing the expression of normal and mutant genes and identifying mutant genes such as oncogenes, in typing tissue for compatibility preceding tissue transplantation, in matching tissue or blood samples for forensic medicine, and for exploring homology among genes from different species.

Ideally, a gene probe assay should be sensitive, specific and easily automatable (for a review, see Nickerson, Current Opinion in Biotechnology 4:48-51 (1993)). The requirement for

sensitivity (i.e. low detection limits) has been greatly alleviated by the development of the polymerase chain reaction (PCR) and other amplification technologies which allow researchers to amplify exponentially a specific nucleic acid sequence before analysis (for a review, see Abramson et al., Current Opinion in Biotechnology, 4:41-47 (1993)).

Specificity, in contrast, remains a problem in many currently available gene probe assays. The extent of molecular complementarity between probe and target defines the specificity of the interaction. Variations in the concentrations of probes, of targets and of salts in the hybridization medium, in the reaction temperature, and in the length of the probe may alter or influence the specificity of the probe/target interaction.

It may be possible under some circumstances to distinguish targets with perfect complementarity from targets with mismatches, although this is generally very difficult using traditional technology, since small variations in the reaction conditions will alter the hybridization. New experimental techniques for mismatch detection with standard probes include DNA ligation assays where single point mismatches prevent ligation and probe digestion assays in which mismatches create sites for probe cleavage.

Recent focus has been on the analysis of the relationship between genetic variation and phenotype by making use of polymorphic DNA markers. Previous work utilized short tandem repeats (STRs) as polymorphic positional markers; however, recent focus is on the use of single nucleotide polymorphisms (SNPs), which occur at an average frequency of more than 1 per kilobase in human genomic DNA. Some SNPs, particularly those in and around coding sequences, are likely to be the direct cause of therapeutically relevant phenotypic variants and/or disease predisposition. There are a number of well known polymorphisms that cause clinically important phenotypes; for example, the apoE2/3/4 variants are associated with different relative risk of Alzheimer's and other diseases (see Cordor et al., Science 261(1993). Multiplex PCR amplification of SNP loci with subsequent hybridization to oligonucleotide arrays has been shown to be an accurate and reliable method of simultaneously genotyping at least hundreds of SNPs; see Wang et al., Science, 280:1077 (1998); see also Schafer et al., Nature Biotechnology 16:33-39 (1998). However, in Wang et al. only 50% of 558 SNPs were amplified successfully in a single multiplexed amplification reaction. As such, there exists a need for methods that increase the fidelity and robustness of multiplexing assays.

Accordingly, highly multiplexed detection or genotyping of nucleic acid sequences is desirable to permit a new scale of genetic analysis. Simultaneously detecting many hundreds, to multiple thousands of nucleic acid sequences, will require methods which are sensitive and specific despite high background complexity. In order for such reactions to be conducted at

low cost to permit widespread use of such techniques, uniform sample preparation and reaction conditions must be applied, preferably in an automatable fashion. A variety of various nucleic acid reaction schemes, amplification techniques, and detection platforms have been used in the past toward this end goal, but none have been able to robustly achieve sensitive, accurate levels of multiplexing beyond a few hundred loci.

Accordingly, it is an object of the invention to provide a very sensitive and accurate multiplexed approach for nucleic acid detection with uniform sample preparation and reaction conditions.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a flow chart for array based detection of gene expression.

Figure 2 depicts a flow chart for array-based detection of RNA Alternative Splicing.

Figure 3 depicts a flow chart for genome-wide expression profiling using oligonucleotide-ligation strategy.

Figure 4 depicts a flow chart for genome-wide RNA alternative splicing monitoring using oligonucleotide-ligation strategy.

Figure 5 depicts a flow chart for direct genotyping using a whole-genome oligonucleotide-ligation strategy.

Figure 6 depicts a flow chart for whole-genome oligonucleotide-ligation strategy.

Figure 7 depicts a preferred embodiment of the invention utilizing a poly(A)-poly(T) capture to remove unhybridized probes and targets. Target sequence **5** comprising a poly(A) sequence **6** is hybridized to target probe**115** comprising a target specific sequence **70**, an adapter seqeuence **20**, an unstream universal priming site **25**, and a downstream universal priming site **26**. The resulting hybridization complex is contacted with a bead **51** comprising a linker **55** and a poly(T) capture probe **61**.

Figure 8 depicts a preferred embodiment of removing non-hybridized target probes, utilizing an OLA format. Target **5** is hybridized to a first ligation probe **100** comprising a first target specific sequence **15**, detection position **10**, an adapter seqeuence **20**, an unstream universal priming site **25**, and an optional label **30**, and a second ligation probe **110** comprising a second target specific sequence **16**, a downstream universal priming site **26**, and a nuclease

inhibitor **35**. After ligation, denaturation of the hybridization complex and addition of an exonuclease, the ligated target probe**115** and the second ligation probe **110** is all that is left. The addition of this to an array (in this embodiment, a bead array comprising substrate **40**, bead **50** with linker **55** and capture probe **60** that is substantially complementary to the adapter sequence **20**), followed by washing away of the second ligation probe **110** results in a detectable complex.

Figure 9 depicts a preferred rolling circle embodiment utilizing two ligation probes. Target **5** is hybridized to a first ligation probe **100** comprising a first target specific sequence **15**, detection position **10**, an adapter seqeuence **20**, an unstream universal priming site **25**, an adapter sequence **20** and a RCA primer sequence **120**, and a second ligation probe **110** comprising a second target specific sequence **16** and a downstream universal priming site **26**. Following ligation, an RCA sequence **130** is added, comprising a first universal primer **27** and a second universal primer **28**. The priming sites hybridize to the primers and ligation occurs, forming a circular probe. The RCA sequence **130** serves as the RCA primer for subsequent amplification. An optional restriction endonuclease site is not shown.

Figure 10 depicts preferred a rolling circle embodiment utilizing a single target probe. Target **5** is hybridized to a target probe **115** comprising a first target specific sequence **15**, detection position **10**, an adapter sequence **20**, an upstream universal priming site **25**, a RCA priming site **140**, optional label sequence **150** and a second target specific sequence **16**. Following ligation, denaturation, and the addition of the RCA primer and extension by a polymerase, amplicons are generated. An optional restriction endonuclease site is not shown.

Figure 11 depicts two configurations of probes for multiplex detection of analytes. A depicts a probe containing an adapter **20**, an upstream priming site **25** and a target-specific portion, i.e. bioactive agent **160** bound to a target analyte **7**. B depicts a probe containing an adapter **20**, an upstream universal priming site **25**, a downstream universal priming site **26** and a target-specific portion, i.e. bioactive agent **160** bound to a target analyte **7**.

Figure 12 depicts a preferred method for multiplex detection of analytes. Probes containing universal priming sequence **25** and adapters that identify the target analyte to be detected **21, 22 and 23**, and target specific portions, i.e. bioactive agents **161, 162 and 163** are contacted with target analytes **201 and 202**. Probes to which target analytes bind are contacted with universal primers **210** and amplification reaction mixture. Amplicons are detected and serve as an indication of the presence of the target analyte.

SUMMARY OF THE INVENTION

In accordance with the embodiments outlined above, the present invention permits highly multiplexed detection of target analytes. The method includes contacting target analytes with a composition comprising an amplification enzyme and first and second target probes. The first and second target probes comprising a first and second bioactive agent, respectively, that specifically bind to the first and second target molecules. The probes also comprise a first and second adapter sequence, respectively, such that the first adapter sequence identifies the first target molecule and the second adapter sequence identifies the second target molecule, and at least a first and second upstream universal priming sequence, respectively. The first and second adapter sequences, wherein no ligation is performed, to form first and second amplicons, respectively, and detecting the first and second amplicons, whereby the first and second target molecules, respectively, are detected.

In addition, the invention provides a method for multiplex detection of a plurality of target molecules comprising contacting a plurality of target molecules with a composition comprising an amplification enzyme and a plurality of target probes, each comprising a bioactive agent, wherein the bioactive agent binds to discrete target molecules an adapter sequence that identifies the discrete target molecule that binds the bioactive agent and at least a first upstream universal primer, amplifying the adapter sequences, wherein no ligation is performed, to form a plurality of amplicons, and detecting the plurality of amplicons, whereby the plurality of target molecules, are detected.

In addition, present invention permits highly multiplexed nucleic acid detection reactions under uniform sample preparation and reaction conditions. That is, preferably the method includes multiplexing from hundreds to thousands of assays simultaneously, more preferably up to tens of thousands of assays simultaneously, most preferably up to millions of assays. The inventive method preferably includes 1) immobilizing the sample nucleic acids to be interrogated (in a preferred embodiment, genomic DNA) on a capture surface, such as a solid phase (in a preferred embodiment, immobilizing the genomic DNA on beads); 2) simultaneously conducting at least a first step of a nucleic acid detection reaction with the captured nucleic acids (in the preferred embodiment, the nucleic acid detection reaction comprises two phases: the first phase involves the exposure of the sample nucleic acids to a set of sequence-specific probe(s), the second phase involves an enzymatic step to assure specificity of the nucleic acid detection reaction. The probes used include at least one appropriate universal amplification priming site); 3) a stringent wash step to reduce the complexity of the multiplexed probe mixture by washing away unhybridized probes; 4) optionally conducting the second phase of the nucleic acid detection reaction step of 2) above (in the case of for example competitive hybridization as the nucleic acid detection reaction, no second phase is required); 5) releasing the probes from the sample nucleic acid; 6)

amplification of the released probes (exponential or linear amplification schemes such as PCR, or Invader™, ESPIA (see WO 01/20035, which is expressly incorporated herein by reference), T7 amplification or the novel amplification method disclosed in Application patent application 60/305118, filed July 12, 2001, which is expressly incorporated herein by reference, using the universal amplification priming site(s) on the probes; and 6) detection and readout of the amplified signals on any detection platform (in a preferred embodiment, the randomly assembled BeadArray™ technology platform).

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to the multiplex preparation and detection of target analytes. In general, the invention involves the use of probes that comprise a number of components. First of all, the probes comprise a bioactive agent (e.g. one of a binding partner pair) that will bind to all or a portion of the target analyte. This bioactive agent may comprise nucleic acid, for example when the target analyte is a target nucleic acid sequence, or, when the target analyte is a protein, for example, the bioactive agent may be nucleic acid (sometimes referred to as an aptamer) or a binding partner such as an antibody or ligand.

The probes further comprise at least one adapter nucleic acid sequence that uniquely identifies the target analyte. That is, there is a unique adapter sequence/target analyte pair for each unique target analyte, although in some cases, adapter sequences may be reused.

In addition, the probes also comprise at least one universal nucleic acid priming sequence that will allow the amplification of the adapter sequence. In some cases, one universal priming sequence can be used, for example when the priming sequence comprises an RNA polymerase priming sequence such as a T7 site. Alternatively, two universal priming sequences can be used, such as standard PCR priming sequences, as long as they flank the adapter sequence, e.g. one priming sequence is 5' to the adapter sequence and one is 3'.

Once the probes have been added to the target analytes to form assay complexes (sometimes referred to herein as hybridization complexes when the target analytes are nucleic acids) generally the unhybridized probes are washed away, using a variety of techniques as outlined herein.

In one embodiment the single nucleotide polymorphisms are detected as outlined herein. This analysis step is followed by amplification as described below.

Amplification proceeds in a number of ways. In general, when an RNA polymerase priming

sequence is used such as a T7 site, the RNA polymerase is added and copies of the adapter sequence are generated. Alternatively, when the amplification reaction is PCR, two primers are added, each of which is substantially complementary (and preferably perfectly complementary) to one of the universal priming sequences. Again, as outlined more fully below, there may be more than one set of universal priming sequences/primers used in a given reaction. In addition, as will be appreciated by those in the art, a number of other amplification reactions can be done, as outlined below.

In an alternative embodiment, the "Universal" primer sequences are designed not to solely serve as PCR primers, but also as a promoter sequence for RNA Polymerase. Thus, the annealed (and/or ligated) target probes can be amplified not only by general PCR, but can also be amplified by in vitro transcription (IVT). The linear amplification produced by IVT should be better at maintaining the relative amounts of the different sequences in the initial template population than the exponential amplification of PCR.

The resulting amplicons can be detected in a wide variety of ways, including the use of biochips (e.g. solid support arrays, including both ordered and random arrays, as outlined herein) liquid arrays, mass spectroscopy analysis, etc., in a variety of formats, including sandwich assays.

In some cases, one or more of the target analytes or probes may be attached to a solid support. For example, the target analytes (for example, genomic DNA sequences) can be attached to beads in a variety of ways. The probe pool is added to form assay complexes (sometimes referred to herein as hybridization complexes when the target analytes are nucleic acids) and unhybridized probes are washed away. The probes are denatured off the target analytes, and then amplified as outlined herein.

Alternatively, solution phase assays may be done, followed by either liquid or solid array detection.

Accordingly, the present invention relates to the multiplex amplification and detection of target analytes in a sample. As used herein, the phrase or "multiplex" or grammatical equivalents refers to the detection, analysis or amplification of more than one target analyte of interest. In a one embodiment multiplex refers to at least 100 different target analytes while at least 500 different target analytes is preferred. More preferred is at least 1000, with more than 5000 particularly preferred and more than 10,000 most preferred. Detection is performed on a variety of platforms. In a preferred embodiment the invention is utilized with adapter sequences that identify the target molecule.

Accordingly, the present invention provides compositions and methods for detecting target analytes including detecting and genotyping specific target nucleic acid sequences in a sample and detecting an quantitating proteins. As will be appreciated by those in the art, the sample solution may comprise any number of things, including, but not limited to, bodily fluids (including, but not limited to, blood, urine, serum, lymph, saliva, anal and vaginal secretions, perspiration and semen, of virtually any organism, with mammalian samples being preferred and human samples being particularly preferred). The sample may comprise individual cells, including primary cells (including bacteria), and cell lines, including, but not limited to, tumor cells of all types (particularly melanoma, myeloid leukemia, carcinomas of the lung, breast, ovaries, colon, kidney, prostate, pancreas and testes), cardiomyocytes, endothelial cells, epithelial cells, lymphocytes (T-cell and B cell) , mast cells, eosinophils, vascular intimal cells, hepatocytes, leukocytes including mononuclear leukocytes, stem cells such as haemopoetic, neural, skin, lung, kidney, liver and myocyte stem cells, osteoclasts, chondrocytes and other connective tissue cells, keratinocytes, melanocytes, liver cells, kidney cells, and adipocytes. Suitable cells also include known research cells, including, but not limited to, Jurkat T cells, NIH3T3 cells, CHO, Cos, 923, HeLa, WI-38, Weri-1, MG-63, etc. See the ATCC cell line catalog, hereby expressly incorporated by reference.

If required, the target analyte is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, sonication, electroporation, etc., with purification and amplification as outlined below occurring as needed, as will be appreciated by those in the art. In addition, the reactions outlined herein may be accomplished in a variety of ways, as will be appreciated by those in the art. Components of the reaction may be added simultaneously, or sequentially, in any order, with preferred embodiments outlined below. In addition, the reaction may include a variety of other reagents which may be included in the assays. These include reagents like salts, buffers, neutral proteins, e.g. albumin, detergents, etc., which may be used to facilitate optimal hybridization and detection, and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used, depending on the sample preparation methods and purity of the target.

In addition, when nucleic acids are to be detected preferred methods utilize cutting or shearing techniques to cut the nucleic acid sample containing the target sequence into a size that will facilitate handling and hybridization to the target, particularly for genomic DNA samples. This may be accomplished by shearing the nucleic acid through mechanical forces (e.g. sonication) or by cleaving the nucleic acid using restriction endonucleases.

In addition, in most embodiments, double stranded target nucleic acids are denatured to

render them single stranded so as to permit hybridization of the primers and other probes of the invention. A preferred embodiment utilizes a thermal step, generally by raising the temperature of the reaction to about 95°C, although pH changes and other techniques may also be used.

The present invention is directed to methods of detecting target analytes. By "target analyte" or "analyte" or grammatical equivalents herein is meant any molecule, compound or particle to be detected. As outlined below, target analytes preferably bind to binding ligands, as is more fully described herein. As will be appreciated by those in the art, a large number of analytes may be detected using the present methods; basically, any target analyte for which a binding ligand, described herein, may be made may be detected using the methods of the invention.

Suitable analytes include organic and inorganic molecules, including biomolecules. They may include environmental pollutants (including pesticides, insecticides, toxins, etc.); a chemical (including solvents, polymers, organic materials, etc.); therapeutic molecules (including therapeutic and abused drugs, antibiotics, etc.); biomolecules (including hormones, cytokines, proteins, lipids, carbohydrates, cellular membrane antigens and receptors (neural, hormonal, nutrient, and cell surface receptors) or their ligands, etc); whole cells (including procaryotic (such as pathogenic bacteria) and eukaryotic cells, including mammalian tumor cells); viruses (including retroviruses, herpesviruses, adenoviruses, lentiviruses, etc.); and spores; etc. Particularly preferred analytes are nucleic acids and proteins (including enzymes, antibodies, antigens, growth factors, cytokines, etc).

In a preferred embodiment, the target analyte is a protein. As will be appreciated by those in the art, there are a large number of possible proteinaceous target analytes that may be detected or evaluated for binding partners using the present invention. Suitable protein target analytes include, but are not limited to, (1) immunoglobulins; (2) enzymes (and other proteins); (3) hormones and cytokines (many of which serve as ligands for cellular receptors); and (4) other proteins. Thus, in one embodiment the method is directed to the detection and/or quantification of protein molecules. In particular, the invention relates to multiplexed detection of proteins. In a preferred embodiment the invention is utilized with adapter sequences that specifically identify the target protein.

In a preferred embodiment, the target analyte is a nucleic acid. By "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, as outlined below, particularly for use with probes or primers, nucleic acid analogs are included that may have alternate backbones, comprising, for example, phosphoramide (Beaucage et al., Tetrahedron 49(10):1925 (1993) and references

therein; Letsinger, J. Org. Chem. 35:3800 (1970); Sprinzl et al., Eur. J. Biochem. 81:579 (1977); Letsinger et al., Nucl. Acids Res. 14:3487 (1986); Sawai et al, Chem. Lett. 805 (1984), Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); and Pauwels et al., Chemica Scripta 26:141 91986)), phosphorothioate (Mag et al., Nucleic Acids Res. 19:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu et al., J. Am. Chem. Soc. 111:2321 (1989), O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, J. Am. Chem. Soc. 114:1895 (1992); Meier et al., Chem. Int. Ed. Engl. 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson et al., Nature 380:207 (1996), all of which are incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcy et al., Proc. Natl. Acad. Sci. USA 92:6097 (1995); non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowshi et al., Angew. Chem. Intl. Ed. English 30:423 (1991); Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); Letsinger et al., Nucleoside & Nucleotide 13:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker et al., Bioorganic & Medicinal Chem. Lett. 4:395 (1994); Jeffs et al., J. Biomolecular NMR 34:17 (1994); Tetrahedron Lett. 37:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins et al., Chem. Soc. Rev. (1995) pp169-176). Several nucleic acid analogs are described in Rawls, C & E News June 2, 1997 page 35. All of these references are hereby expressly incorporated by reference. These modifications of the ribose-phosphate backbone may be done to facilitate the addition of labels, or to increase the stability and half-life of such molecules in physiological environments.

In a preferred embodiment, the nucleic acid preferably includes at least one universal base. Universal bases are those that can substitute for any of the five natural bases, that is, universal bases will basepair with all natural bases, preferably equally well. Suitable universal bases include, but are not limited to, inosine, hypoxanthine, 5-nitroindole, acylic 5-nitroindole, 4-nitropyrazole, 4-nitroimidazole and 3-nitropyrrole. See Loakes et al., Nucleic Acid Res. 22:4039 (1994); Van Aerschot et al., Nucleic Acid Res. 23:4363 (1995); Nichols et al., Nature 369:492 (1994); Berstrom et al., Nucleic Acid Res. 25:1935 (1997); Loakes et al., Nucleic Acid Res. 23:2361 (1995); Loakes et al., J. Mol. Biol. 270:426 (1997); and Fotin et al., Nucleic Acid Res. 26:1515 (1998); and references cited therein, all of which are expressly incorporated by reference.

As will be appreciated by those in the art, all of these nucleic acid analogs may find use in the

present invention. In addition, mixtures of naturally occurring nucleic acids and analogs can be made. Alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

Particularly preferred are peptide nucleic acids (PNA) which includes peptide nucleic acid analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring nucleic acids. This results in two advantages. First, the PNA backbone exhibits improved hybridization kinetics. PNAs have larger changes in the melting temperature (Tm) for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4°C drop in Tm for an internal mismatch. With the non-ionic PNA backbone, the drop is closer to 7-9°C. This allows for better detection of mismatches. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration.

The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. Thus, for example, when the target sequence is a polyadenylated mRNA, the hybridization complex comprising the target probe has a double stranded portion, where the target probe is hybridized, and one or more single stranded portions, including the poly(A) portion. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xathanine hypoxathanine, isocytosine, isoguanine, etc. A preferred embodiment utilizes isocytosine and isoguanine in nucleic acids designed to be complementary to other probes, rather than target sequences, as this reduces non-specific hybridization, as is generally described in U.S. Patent No. 5,681,702. As used herein, the term "nucleoside" includes nucleotides as well as nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, "nucleoside" includes non-naturally occuring analog structures. Thus for example the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

In a preferred embodiment, the compositions and methods of the invention are directed to the detection of target sequences. The term "target sequence" or "target nucleic acid" or grammatical equivalents herein means a nucleic acid sequence on a single strand of nucleic acid. The target sequence may be a portion of a gene, a regulatory sequence, genomic DNA (gDNA), cDNA, RNA including mRNA and rRNA, or others, with polyadenylated mRNA being particular preferred in some embodiments. As is outlined herein, the target sequence may be a target sequence from a sample, or a secondary target such as an amplicon, which is the product of an amplification reaction such as PCR or an RNA polymerase reaction. Thus, for example, a target sequence from a sample is amplified to produce an amplicon that is

detected. The target sequence may be any length, with the understanding that longer sequences are more specific. As will be appreciated by those in the art, the complementary target sequence may take many forms. For example, it may be contained within a larger nucleic acid sequence, i.e. all or part of a gene or mRNA, a restriction fragment of a plasmid or genomic DNA, among others. Particularly preferred target sequences in the present invention include genomic DNA, polyadenylated mRNA, and alternatively spliced RNAs. As is outlined more fully below, probes are made to hybridize to target sequences to determine the presence, absence, quantity or sequence of a target sequence in a sample. Generally speaking, this term will be understood by those skilled in the art.

The target sequence may also be comprised of different target domains, that may be adjacent (i.e. contiguous) or separated. The terms "first" and "second" are not meant to confer an orientation of the sequences with respect to the 5'-3' orientation of the target sequence. For example, assuming a 5'-3' orientation of the complementary target sequence, the first target domain may be located either 5' to the second domain, or 3' to the second domain. In addition, as will be appreciated by those in the art, the probes on the surface of the array (e.g. attached to the microspheres) may be attached in either orientation, either such that they have a free 3' end or a free 5' end; in some embodiments, the probes can be attached at one ore more internal positions, or at both ends.

In a preferred embodiment the invention is directed to target sequences that comprise one or more positions for which sequence information is desired, generally referred to herein as the "detection position" or "detection locus". In a preferred embodiment, the detection position is a single nucleotide (sometimes referred to as a single nucleotide polymorphism (SNP)), although in some embodiments, it may comprise a plurality of nucleotides, either contiguous with each other or separated by one or more nucleotides. By "plurality" as used herein is meant at least two. As used herein, the base of a probe (e.g. the target probe) which basepairs with a detection position base in a hybrid is termed a "readout position" or an "interrogation position". Thus, the target sequence comprises a detection position and the target probe comprises a readout position.

In a preferred embodiment, the use of competitive hybridization target probes is done to elucidate either the identity of the nucleotide(s) at the detection position or the presence of a mismatch.

It should be noted in this context that "mismatch" is a relative term and meant to indicate a difference in the identity of a base at a particular position, termed the "detection position" herein, between two sequences. In general, sequences that differ from wild type sequences are referred to as mismatches. However, particularly in the case of SNPs, what constitutes

"wild type" may be difficult to determine as multiple alleles can be relatively frequently observed in the population, and thus "mismatch" in this context requires the artificial adoption of one sequence as a standard. Thus, for the purposes of this invention, sequences are referred to herein as "match" and "mismatch". Thus, the present invention may be used to detect substitutions, insertions or deletions as compared to a wild-type sequence. That is, all other parameters being equal, a perfectly complementary readout target probe (a "match probe") will in general be more stable and have a slower off rate than a target probe comprising a mismatch (a "mismatch probe") at any particular temperature.

In some embodiments, as outlined below, the target analytes (or target probes, in some instances) may be attached to a solid support prior to contact with the target probes (or to remove unhybridized target probes, etc.). In this embodiment, the target analyte may comprise a purification tag. By "purification tag" herein is meant a moiety which can be used to purify a strand of nucleic acid, usually via attachment to a solid support as outlined herein. Suitable purification tags include members of binding partner pairs. For example, the tag may be a hapten or antigen, which will bind its binding partner. In a preferred embodiment, the binding partner can be attached to a solid support as depicted herein and in the figures. For example, suitable binding partner pairs include, but are not limited to: antigens (such as proteins (including peptides)) and antibodies (including fragments thereof (FAbs, etc.)); proteins and small molecules, including biotin/streptavidin; enzymes and substrates or inhibitors; other protein-protein interacting pairs; receptor-ligands; and carbohydrates and their binding partners. Nucleic acid - nucleic acid binding proteins pairs are also useful. In general, the smaller of the pair is attached to the NTP for incorporation into the primer. Preferred binding partner pairs include, but are not limited to, biotin (or imino-biotin) and streptavidin, digeoxinin and Abs, and Prolinx™ reagents (see www.prolinxinc.com/ie4/home.hmtl).

In a preferred embodiment, the binding partner pair comprises biotin or imino-biotin and streptavidin. Imino-biotin is particularly preferred as imino-biotin disassociates from streptavidin in pH 4.0 buffer while biotin requires harsh denaturants (e.g. 6 M guanidinium HCl, pH 1.5 or 90% formamide at 95°C).

The present invention provides methods and compositions directed to the multiplex amplification and detection of target sequences utilizing target probes.

Accordingly, the invention provides a number of different primers and probes. Many of the probes and primers of the present invention are designed to have at least a portion that binds substantially specifically to a target analyte (sometimes referred to herein as a bioactive agent (particularly in the case wherein the target analyte is not a nucleic acid) or a target specific portion). That is the probes are constructed so as to contain a target specific portion: a

portion that binds to the target analyte specifically, i.e. with high affinity. This target specific portion can be any type of molecule so long as it specifically binds the target and can be attached to the rest of a target probe, namely a nucleic acid sequence that preferably includes an adapter sequence and at least one priming sequence.

In a preferred embodiment, when the target analyte is a protein, for example, the target specific portion ("bioactive agent") may include other protein molecules such as antibodies, specific binding proteins or ligands such as growth factors. Also, the target-specific portion can be an aptamer as is known in the art, i.e. a nucleotide sequence that binds with high affinity to a protein.

In a preferred embodiment, the binding of the bioactive agent and the target analyte is specific; that is, the bioactive agent specifically binds to the target analyte. By "specifically bind" herein is meant that the agent binds the analyte, with specificity sufficient to differentiate between the analyte and other components or contaminants of the test sample. However, as will be appreciated by those in the art, it will be possible to detect analytes using binding which is not highly specific; for example, the systems may use different binding ligands, for example an array of different ligands, and detection of any particular analyte is via its "signature" of binding to a panel of binding ligands, similar to the manner in which "electronic noses" work. This finds particular utility in the detection of chemical analytes. The binding should be sufficient to remain bound under the conditions of the assay, including wash steps to remove non-specific binding, although in some embodiments, wash steps are not desired; i.e. for detecting low affinity binding partners. In some embodiments, for example in the detection of certain biomolecules, the dissociation constants of the analyte to the binding ligand will be less than about 10-4-10-6 M-1, with less than about 10-5 to 10-9 M-1 being preferred and less than about 10-7 -10-9 M-1 being particularly preferred.

When nucleic acids are the target, the probes are designed to be complementary to all or a portion (domain) of a target sequence (either the target sequence of the sample or to other probe sequences, such as portions of amplicons, as is described below), such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by "substantially complementary" herein is meant that the bioactive agent portion of the probes are sufficiently complementary to all or part of the target sequences to hybridize under normal reaction conditions, and preferably give the required specificity. In a preferred embodiment the probes

have a portion that is exactly complementary to the target nucleic acids.

A variety of hybridization conditions may be used in the present invention, including high, moderate and low stringency conditions; see for example Maniatis et al., Molecular Cloning: A Laboratory Manual, 2d Edition, 1989, and Short Protocols in Molecular Biology, ed. Ausubel, et al, hereby incorporated by reference. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at Tm, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of helix destabilizing agents such as formamide. The hybridization conditions may also vary when a non-ionic backbone, i.e. PNA is used, as is known in the art. In addition, cross-linking agents may be added after target binding to cross-link, i.e. covalently attach, the two strands of the hybridization complex.

In a preferred embodiment, the target probes further comprise one or more "adapter sequences" (sometimes referred to in the art as "zip codes") to allow the use of "universal arrays". That is, arrays are generated that contain capture probes that are not target specific, but rather specific to individual artificial adapter sequences. The adapter sequences are added to the target probes, nested between the priming sequences (when two priming sequences are used) or "downstream" of a single universal priming sequence, and thus are included in the amplicons. What is important is that the orientation of the priming sequence and the adapter sequence allows the amplification of the adapter sequence.

An "adapter sequence" is a sequence, generally exogeneous to the target sequences, e.g. artificial, that is designed to be substantially complementary (and preferably perfectly complementary) to a capture probe of a detection array. Generally the capture probe is immobilized to a solid support that can include microspheres or planar substrates such as plastic or glass slides as described herein for array supports. In one embodiment the use of adapter sequences allow the creation of more "universal" surfaces; that is, one standard array,

comprising a finite set of capture probes can be made and used in any application. The end-user can customize the array by designing different soluble target probes, which, as will be appreciated by those in the art, is generally simpler and less costly. In a preferred embodiment, an array of different and usually artificial capture probes are made; that is, the capture probes do not have complementarity to known target sequences. The adapter sequences can then be incorporated in the target probes.

As will be appreciated by those in the art, the length of the adapter sequences will vary, depending on the desired "strength" of binding and the number of different adapters desired. In a preferred embodiment, adapter sequences range from about 6 to about 500 basepairs in length, with from about 8 to about 100 being preferred, and from about 10 to about 25 being particularly preferred.

In a preferred embodiment, the adapter sequence uniquely identifies the target analyte to which the target probe binds. That is, while the adapter sequence need not bind itself to the target analyte, the system allows for identification of the target analyte by detecting the presence of the adapter. Accordingly, following a binding or hybridization assay and washing, the probes including the adapters are amplified. Detection of the adapter then serves as an indication of the presence of the target analyte.

In one embodiment the adapter includes both an identifier region and a region that is complementary to capture probes on a universal array as described above. In this embodiment, the amplicon hybridizes to capture probes on a universal array. Detection of the adapter is accomplished following hybridization with a probe that is complementary to the adapter sequence. Preferably the probe is labeled as described herein.

In general, unique adapter sequences are used for each unique target analyte. That is, the elucidation or detection of a particular adapter sequence allows the identification of the target analyte to which the target probe containing that adapter sequence bound. However, in some cases, it is possible to "reuse" adapter sequences and have more than one target analyte share an adapter sequence.

In a preferred embodiment the adapters contain different sequences or properties that are indicative of a particular target molecule. That is, each adapter uniquely identifies a target analyte. As described above, the adapters are amplified to form amplicons. The adapter is detected as an indication of the presence of the target analyte.

The use of adapters in combination with amplification following a specific binding event allows for highly multiplexed reactions to be performed.

Also, the probes are constructed so as to contain the necessary priming site or sites for the subsequent amplification scheme. In a preferred embodiment the priming sites are universal priming sites. By "universal priming site" or "universal priming sequences" herein is meant a sequence of the probe that will bind a primer for amplification.

In a preferred embodiment, one universal priming sequence or site is used. In this embodiment, a preferred universal priming sequence is the RNA polymerase T7 sequence, that allows the T7 RNA polymerase make RNA copies of the adapter sequence as outlined below.

In a preferred embodiment, for example when amplification methods requiring two primers such as PCR are used, each probe preferably comprises an upstream universal priming site (UUP) and a downstream universal priming site (DUP). Again, "upstream" and "downstream" are not meant to convey a particular 5'- 3' orientation, and will depend on the orientation of the system. Preferably, only a single UUP sequence and a single DUP sequence is used in a probe set, although as will be appreciated by those in the art, different assays or different multiplexing analysis may utilize a plurality of universal priming sequences. In addition, the universal priming sites are preferably located at the 5' and 3' termini of the target probe (or the ligated probe), as only sequences flanked by priming sequences will be amplified.

In addition, universal priming sequences are generally chosen to be as unique as possible given the particular assays and host genomes to ensure specificity of the assay. However, as will be appreciated by those in the art, sets of priming sequences/primers may be used; that is, one reaction may utilize 500 target probes with a first priming sequence or set of sequences, and an additional 500 probes with a second sequence or set of sequences.

As will be appreciated by those in the art, when two priming sequences are used, the orientation of the two priming sites is different. That is, one PCR primer will directly hybridize to the first priming site, while the other PCR primer will hybridize to the complement of the second priming site. Stated differently, the first priming site is in sense orientation, and the second priming site is in antisense orientation.

The size of the primer and probe nucleic acid may vary, as will be appreciated by those in the art with each portion of the probe and the total length of the probe in general varying from 5 to 500 nucleotides in length. Each portion is preferably between 10 and 100 being preferred, between 15 and 50 being particularly preferred, and from 10 to 35 being especially preferred, depending on the use and amplification technique. Thus, for example, the universal priming site(s) of the probes are each preferably about 15-20 nucleotides in length, with 18 being especially preferred. The adapter sequences of the probes are preferably from 15-25

nucleotides in length, with 20 being especially preferred. The target specific portion of the probe is preferably from 15-50 nucleotides in length. In addition, the primer may include an additional amplification priming site. In a preferred embodiment the additional amplification priming site is a T7 RNA polymerase priming site.

Accordingly, the present invention provides first target probe sets. By "probe set" herein is meant a plurality of target probes that are used in a particular multiplexed assay. In this context, plurality means at least two, with more than 10 being preferred, depending on the assay, sample and purpose of the test. In one embodiment the probe set includes more than 100, with more than 500 probes being preferred and more than 1000 being particularly preferred. In a particularly preferred embodiment each probe contains at least 5000, with more than 10,000 probes being most preferred.

Accordingly, the present invention provides first target probe sets that comprise at least a first universal priming site.

In a preferred embodiment, the target probe may also comprise a label sequence, i.e. a sequence that can be used to bind label probes and is substantially complementary to a label probe. This is sometimes referred to in the art as "sandwich-type" assays. That is, by incorporating a label sequence into the target probe, which is then amplified and present in the amplicons, a label probe comprising primary (or secondary) labels can be added to the mixture, either before addition to the array or after. This allows the use of high concentrations of label probes for efficient hybridization. In one embodiment, it is possible to use the same label sequence and label probe for all target probes on an array; alternatively, different target probes can have a different label sequence. Similarly, the use of different label sequences can facilitate quality control; for example, one label sequence (and one color) can be used for one strand of the target, and a different label sequence (with a different color) for the other; only if both colors are present at the same basic level is a positive called.

Thus, the present invention provides target probes that comprise universal priming sequences, bioactive agents (e.g. target specific portion(s)), adapter sequence(s), optionally an additional amplification priming sequence such as T7 RNA priming sequence and optionally label sequences. These target probes are then added to the target sequences to form hybridization complexes. As will be appreciated by those in the art, when nucleic acids are the target, the hybridization complexes contain portions that are double stranded (the target-specific sequences of the target probes hybridized to a portion of the target sequence) and portions that are single stranded (the ends of the target probes comprising the universal priming sequences and the adapter sequences, and any unhybridized portion of the target sequence, such as poly(A) tails, as outlined herein).

As will be appreciated by those in the art, the systems of the invention can take on a wide variety of configurations, including systems that rely on the initial immobilization of the target analyte (solid phase assays) and solution based assays.

Solid phase assays

In a preferred embodiment, the target analyte is immobilized on the surface. That is, the target analytes including proteins or target nucleic acids or target sequences are immobilized on a substrate or capture surface. Attachment may be performed in a variety of ways, as will be appreciated by those in the art, including, but not limited to, chemical or affinity capture (for example, including the incorporation attachment moieties such as derivatized nucleotides such as AminoLink™ or biotinylated nucleotides that can then be used to attach the nucleic acid to a surface, as well as affinity capture by hybridization), cross-linking, and electrostatic attachment, etc. When the target analyte is polyadenylated mRNA, supports comprising poly(T) sequences can be used. That is, an attachment moiety is attached to the target analyte that allows for attachment to the substrate. By "attachment moiety" is meant a molecule or substance that mediates attachment of the target analyte to the substrate. In a preferred embodiment, affinity capture is used to attach the nucleic acids to the support. For example, nucleic acids or proteinscan be derivatized, for example with one member of a binding pair, and the support derivatized with the other member, i.e. a complementary member, of a binding pair. For example, the nucleic acids may be biotinylated (for example using enzymatic incorporation of biotinylated nucleotides, or by photoactivated cross-linking of biotin). In a preferred embodiment the target nucleic acids are photobiotinylated In one preferred embodiment the target nucleic acids are photobiotinylated with PHOTOPROBE™ Biotin Reagents (Vector Laboratories). Biotinylated nucleic acids can then be captured on streptavidin-coated surfaces, as is known in the art. In one embodiment the surfaces or supports are beads to which the nucleic acids are attached. In a particularly preferred embodiment the beads are magnetic beads. Similarly, other hapten-receptor combinations can be used, such as digoxigenin and anti-digoxigenin antibodies. Alternatively, chemical groups can be added in the form of derivatized nucleotides, that can then be used to add the nucleic acid to the surface.

Similarly, affinity capture utilizing hybridization can be used to attach nucleic acids to surface or bead.

For example, a polyA tract can be attached by polymerization with terminal transferase, or via ligation of an oligoA linker, as is known in the art. This then allows for hybridization with an immobilized poly-T tract. Alternatively, chemical crosslinking may be done, for example by photoactivated crosslinking of thymidine to reactive groups, as is known in the art.

Preferred attachments are covalent, although even relatively weak interactions (i.e. non-covalent) can be sufficient to attach a nucleic acid to a surface, if there are multiple sites of attachment per each nucleic acid. Thus, for example, electrostatic interactions can be used for attachment, for example by having beads carrying the opposite charge to the bioactive agent.

A preferred embodiment utilizes covalent attachment of the target sequences to a support. As is known in the art, there are a wide variety of methods used to covalently attach nucleic acids to surfaces. A preferred embodiment utilizes the incorporation of a chemical functional group into the nucleic acid, followed by reaction with a derivatized or activated surface. Examples include, but are not limited to AminoLink™.

By "capture surface", "target substrate" or "target support" or other grammatical equivalents herein is meant any material to which a target analyte can be attached. The targets can be attached either directly or indirectly as described herein. As will be appreciated by those in the art, the number of possible substrates is very large. Possible substrates include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, and a variety of other polymers. Preferably the substrates include microfuge tubes, i.e. Eppendorf tubes. In one embodiment the substrates include beads or microspheres. In one embodiment the beads or microspheres are magnetic. In one embodiment the substrates are derivatized to accommodate attachment of the target nucleic acids to the substrate.

The configuration of the target support is not crucial. What is important is that the target analytes are immobilized to the target support and can be manipulated. That is, the support should be amenable to a variety of reactions as described herein. While the target substrate can be flat (planar), other configurations of substrates may be used as well; for example, target analytes can be attached to beads or microspheres that can be deposited in reaction tubes or vessels or wells. That is, the target substrate may be microspheres to which the target analytes are attached. The microspheres can then be distributed on a surface. In some embodiments the surface contains reaction wells into which the beads are distributed, for example microtiter plates as are known in the art and as described herein.

Once the target analytes, i.e. genomic DNA or proteins, are applied to or immobilized on the surface, the target analytes are contacted with probes for analyses, including detection or genotyping. That is, the appropriate probes necessary for detection of the target analyte or for

the mutation detection reactions are next introduced to the immobilized sample.

For the assays described herein, the assays may be run under a variety of experimental conditions, as will be appreciated by those in the art. A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding. Various blocking and washing steps may be utilized as is known in the art.

Following binding or hybridization of the bioactive agent portion of the target probe to the target analyte, unhybridized probes are removed by a washing step. In a preferred embodiment the wash step is a stringent wash step. That is, in the preferred embodiment of an enzymatic based mutation detection reaction, once the probes have been introduced under conditions to favor hybridization with the appropriate nucleic acid sequences, a stringent wash step is conducted. This wash removes unhybridized probes and reduces the overall complexity of the mixture. It is this step that ensures the success of the overall multiplexed reaction.

As will be appreciated by those in the art, when the target analyte is a nucleic acid, there are a wide variety of SNP detection reactions that can be done at this stage. In a preferred embodiment, different target probes are made with different interrogation bases in the target-specific domain of the probe. The wash step is done under conditions to wash away imperfect matches between the interrogation base of the target probe and the detection position of the target sequence.

In a preferred embodiment when nucleic acids are the target, a plurality of target probes (sometimes referred to herein as "readout target probes") are used to identify the base at the detection position. In this embodiment, each different readout probe comprises a different base at the position that will hybridize to the detection position of the target sequence (herein referred to as the readout or interrogation position) and a different adapter sequence for each different readout position. In this way, differential hybridization of the readout target probes, depending on the sequence of the target, results in identification of the base at the detection position. In this embodiment, the readout probes are contacted with the array again under conditions that allow discrimination between match and mismatch, and the unhybridized probes are removed, etc.

Accordingly, by using different readout target probes, each with a different base at the readout position and each with a different adapter, the identification of the base at the detection position is elucidated. Thus, in a preferred embodiment, a set of readout probes are used, each comprising a different base at the readout position.

In a preferred embodiment, each readout target probe has a different adapter sequence. That is, readout target probes comprising adenine at the readout position will have a first adapter, probes with guanine at the readout position will have a second adapter, etc., such that each target probe that hybridizes to the target sequence will bind to a different address on the array. This can allow the use of the same label for each reaction.

The number of readout target probes used will vary depending on the end use of the assay. For example, many SNPs are biallelic, and thus two readout target probes, each comprising an interrogation base that will basepair with one of the detection position bases. For sequencing, for example, for the discovery of SNPs, a set of four readout probes are used.

In this embodiment, sensitivity to variations in stringency parameters are used to determine either the identity of the nucleotide(s) at the detection position or the presence of a mismatch. As a preliminary matter, the use of different stringency conditions such as variations in temperature and buffer composition to determine the presence or absence of mismatches in double stranded hybrids comprising a single stranded target sequence and a probe is well known.

With particular regard to temperature, as is known in the art, differences in the number of hydrogen bonds as a function of basepairing between perfect matches and mismatches can be exploited as a result of their different Tms (the temperature at which 50% of the hybrid is denatured). Accordingly, a hybrid comprising perfect complementarity will melt at a higher temperature than one comprising at least one mismatch, all other parameters being equal. (It should be noted that for the purposes of the discussion herein, all other parameters (i.e. length of the hybrid, nature of the backbone (i.e. naturally occuring or nucleic acid analog), the assay solution composition and the composition of the bases, including G-C content are kept constant). However, as will be appreciated by those in the art, these factors may be varied as well, and then taken into account.)

In general, as outlined herein, high stringency conditions are those that result in perfect matches remaining in hybridization complexes, while imperfect matches melt off. Similarly, low stringency conditions are those that allow the formation of hybridization complexes with both perfect and imperfect matches. High stringency conditions are known in the art as outlined above.

As will be appreciated by those in the art, mismatch detection using temperature may proceed in a variety of ways.

Similarly, variations in buffer composition may be used to elucidate the presence or absence of a mismatch at the detection position. Suitable conditions include, but are not limited to, formamide concentration. Thus, for example, "low" or "permissive" stringency conditions include formamide concentrations of 0 to 10%, while "high" or "stringent" conditions utilize formamide concentrations of ≥40%. Low stringency conditions include NaCl concentrations of ≥1 M, and high stringency conditions include concentrations of ≤ 0.3 M. Furthermore, low stringency conditions include $MgCl_2$ concentrations of ≥ 10 mM, moderate stringency as 1-10 mM, and high stringency conditions include concentrations of ≤ 1 mM.

In this embodiment, as for temperature, a plurality of readout probes may be used, with different bases in the readout position and different adapters. Running the assays under the permissive conditions and repeating under stringent conditions will allow the elucidation of the base at the detection position.

Thus, the washing is performed under stringency conditions which allows formation of the first hybridization complex only between probes and complementary target sequences. As outlined above, stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration, pH, organic solvent concentration, etc.

These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Patent No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

In a preferred embodiment, the target sequence may be immobilized after the formation of the hybridization complexes, ligation complexes and/or ligated complexes. That is, the probes can be added to the targets in solution, enzymes added as needed, etc. After the hybridization complexes are formed and/or ligated, the hybridization complexes can be added to supports comprising the binding partners and the unhybridized probes removed.

In this embodiment, particularly preferred binding ligand/binding partner pairs are biotin and streptavidin or avidin, antigens and antibodies.

As described above, once the hybridization complexes are formed, unhybridized probes are removed. This is important to increase the level of multiplexing in the assay. In addition, as all target probes may form some unpredictable structures that will complicate the amplification

using the universal priming sequences. Thus to ensure specificity (e.g. that target probes directed to target sequences that are not present in the sample are not amplified and detected), it is important to remove all the nonhybridized probes. As will be appreciated by those in the art, this may be done in a variety of ways, including methods based on the target sequence, methods utilizing double stranded specific moieties, and methods based on probe design and content. Preferably the method includes a stringent wash step.

Once the non-hybridized probes (and additionally, if preferred, other sequences from the sample that are not of interest) are removed, the hybridization complexes are denatured and the target probes are amplified to form amplicons, which are then detected. This can be done in one of several ways as outlined below. In addition, as outlined below, labels can be incorporated into the amplicons in a variety of ways.

Accordingly, this embodiment can be run in several modes. In a preferred embodiment, only a single probe is used, comprising (as outlined herein), at least a first UUP, an adapter sequence, and a target-specific portion, i.e. a target specific moeity or bioactive agent. When nucleotides are the target molecule the target-specific portion includes nucleic acids comprising a first base at the readout position, and in some embodiments a DUP. This probe is contacted with the target analyte under conditions (whether thermal or otherwise) such that specific binding occurs. In a preferred embodiment, when nucleic acids are the target, a hybridization complex is formed only when a perfect match between the detection position of the target and the readout position of the probe is present. The non-hybridized or non-bound probes are then removed as outlined herein. That is, after the wash step, only the properly hybridized probes should remain. In one embodiment when nucleic acids are the target, the hybridized probes must then be separated from the captured sample nucleic acid. This is done via a stringent wash step or denaturation step. The sample nucleic acid is left behind on the capture surface, and can be used again. In an alternative embodiment, although not preferred, the hybridized probe is not removed. it is not necessary to remove the probes when the priming sites and adapter sequences do not hybridize with the target. The probe is then amplified as outlined herein, and detected. In a preferred embodiment the amplified product(s), i.e. amplicons, are detected as an indication of the presence of the target analyte.

When molecules besides nucleic acids are the target, the bound probes need not be removed from the binding complex prior to amplification. That is, amplification can proceed while the probes are bound to the target. Without being bound be theory, it is thought that amplification can proceed because the nucleotides are not hybridized with a complementary strand. As such, they are free to hybridize with amplification primers.

As noted above, the target sequence may be immobilized either before or after the formation

of the hybridization complex, but preferably it is immobilized on a surface or support comprising the binding partner of the binding ligand prior to the formation of the hybridization complex with the probe(s) of the invention. For example, a preferred embodiment utilizes binding partner coated reaction vessels such as eppendorf tubes or microtiter wells. Alternatively, the support may be in the form of beads, including magnetic beads. In this embodiment, the target sequences are immobilized, the target probes are added to form hybridization complexes. Unhybridized probes are then removed through washing steps, and the bound probes (e.g. either target probes, ligated probes, or ligated RCA probes) are then eluted off the support, usually through the use of elevated temperature or buffer conditions (pH, salt, etc.).

Once the non-hybridized probes (and additionally, if preferred, other sequences from the sample that are not of interest) are removed, the hybridization complexes are denatured and the target probes are amplified to form amplicons, which are then detected. This can be done in one of several ways, including PCR amplification and rolling circle amplification. Also, the probes can be amplified by known methods (exponential or linear amplification techniques such as PCR, Invader, ESPIA (also known as SPIA), T7), using the one or more priming sites provided on the probes. As noted herein, the probes are constructed so as to contain the necessary primer sites to permit this amplification. In a preferred embodiment, universal primers are used. Amplification provides the signal strength and dynamic range necessary for detection of the mutation-detection probes. In addition, as outlined below, labels can be incorporated into the amplicons in a variety of ways.

In a preferred embodiment, no ligation assay for genotyping is done, that is, no ligase is added. However, as will be appreciated by those in the art, ligation reactions for other purposes may be done.

In a preferred embodiment, a linear amplification scheme known as ESPIA, or SPIA is applied. This amplification technique is disclosed in WO 01/20035 A2 and U.S. Application Serial Number 6,251,639, which are incorporated by reference herein. Generally, the method includes hybridizing chimeric RNA/DNA amplification primers to the probes. Preferably the DNA portion of the probe is 3' to the RNA. Optionally the method includes hybridizing a polynucleotide comprising a termination polynucleotide sequence to a region of the template that is 5' with respect to hybridization of the composite primer to the template. Following hybridization of the primer to the template, the primer is extended with DNA polymerase. Subsequently, the RNA is cleaved from the composite primer with an enzyme that cleaves RNA from an RNA/DNA hybrid. Subsequently, an additional RNA/DNA chimeric primer is hybridized to the template such that the first extended primer is displaced from the target probe. The extension reaction is repeated, whereby multiple copies of the probe sequence

are generated.

In a preferred embodiment, the target amplification technique is PCR. The polymerase chain reaction (PCR) is widely used and described, and involves the use of primer extension combined with thermal cycling to amplify a target sequence; see U.S. Patent Nos. 4,683,195 and 4,683,202, and PCR Essential Data, J. W. Wiley & sons, Ed. C.R. Newton, 1995, all of which are incorporated by reference.

In general, PCR may be briefly described as follows. The double stranded hybridization complex is denatured, generally by raising the temperature, and then cooled in the presence of an excess of a PCR primer, which then hybridizes to the first universal priming site. A DNA polymerase then acts to extend the primer with dNTPs, resulting in the synthesis of a new strand forming a hybridization complex. The sample is then heated again, to disassociate the hybridization complex, and the process is repeated. By using a second PCR primer for the complementary target strand that hybridizes to the second universal priming site, rapid and exponential amplification occurs. Thus PCR steps are denaturation, annealing and extension. The particulars of PCR are well known, and include the use of a thermostable polymerase such as Taq I polymerase and thermal cycling. Suitable DNA polymerases include, but are not limited to, the Klenow fragment of DNA polymerase I, SEQUENASE 1.0 and SEQUENASE 2.0 (U.S. Biochemical), T5 DNA polymerase and Phi29 DNA polymerase.

The reaction is initiated by introducing the target probe comprising the target sequence to a solution comprising the universal primers, a polymerase and a set of nucleotides. By "nucleotide" in this context herein is meant a deoxynucleoside-triphosphate (also called deoxynucleotides or dNTPs, e.g. dATP, dTTP, dCTP and dGTP). In some embodiments, as outlined below, one or more of the nucleotides may comprise a detectable label, which may be either a primary or a secondary label. In addition, the nucleotides may be nucleotide analogs, depending on the configuration of the system. Similarly, the primers may comprise a primary or secondary label.

Accordingly, the PCR reaction requires at least one PCR primer, a polymerase, and a set of dNTPs. As outlined herein, the primers may comprise the label, or one or more of the dNTPs may comprise a label.

In a preferred embodiment, instead of using two primers (e.g.unlabeled T3 and biotin-labeled T7), a third primer (overlapping with T7, but is shorter than T7; labeled with another dye, for example, Fam) is added to the PCR reaction. The PCR is first carried out at a lower stringent condition for a certain cycles, i.e. 25-30 cycles, in which both the longer and shorter PCR primers are annealed to the targets and generate PCR products; The PCR is then carried out

at a higher stringent condition for additional cycles, say additional 5-10 cycles. Under this higher stringent condition, only the longer PCR primer can anneal to the targets and further generate PCR products, while the shorter PCR primer will not hybridize under this condition. Accordingly, for each of the target, two PCR products are generated with different PCR cycles and labeled with different dyes. Since the two products are presented at different concentrations in the final hybridization solution, the "shorter primer" signal can be used to measure the genes expressed at high level without running into saturation problem, while the "longer primer" signal is used to measure the genes expressed at low level without losing the sensitivity. While the invention is described using two primer variants, i.e. long and short prove, more than two variants can be used. That is, preferably more than two primer variants are used with more than five being particularly preferred.

In addition, identical primers can be used, but the primers bear different labels. In this embodiment the ratio of the two labels in the product can be adjusted by varying the initial primer concentrations, so there is no need to vary the PCR conditions.

In an alternative embodiment amplification can be performed using two or more dye labeled dNTP (for the PCR) or NTP (for the IVT), pre-mixed at different ratios. Accordingly, there is no need to vary the PCR conditions and PCR primer labeling. This method can also be used in the IVT step in gene expression monitoring using a direct hybridization with total RNA or mRNA, as a way to control the signal saturation problem. As such, detection of labels of different intensity serves to increase the range of detection of targets. That is, using less intense labels allows for detection of abundant targets without saturation while the use of stronger labels serves to increase sensitivity allowing for detection of less abundant targets.

In addition, the methods described above can be used in the final PCR step in OLA-PCR genotyping as well, as long as the dyes are chosen correctly such that they can be well-resolved by the hardware and/or software of the systems. That is following the OLA reaction, the ligation products can be amplified using primers as described above, i.e. either primer variants or differently labeled primers.

In a preferred embodiment, the methods of the invention include a rolling circle amplification (RCA) step. This may be done in several ways. In one embodiment, either single target probes or ligated probes can be used in the genotyping part of the assay, followed by RCA instead of PCR. Alternatively, and more preferably, the RCA reaction forms part of the genotyping reaction and can be used for both genotyping and amplification in the methods of the reaction.

In a preferred embodiment, the methods rely on rolling circle amplification. "Rolling circle

amplification" is based on extension of a circular probe that has hybridized to a target sequence. A polymerase is added that extends the probe sequence. As the circular probe has no terminus, the polymerase repeatedly extends the circular probe resulting in concatamers of the circular probe. As such, the probe is amplified. Rolling-circle amplification is generally described in Baner *et al.* (1998) *Nuc. Acids Res.* 26:5073-5078; Barany, F. (1991) *Proc. Natl. Acad. Sci. USA* 88:189-193; and Lizardi *et al.* (1998) *Nat. Genet.* 19:225-232, all of which are incorporated by reference in their entirety.

In general, RCA may be described in two ways, as generally depicted in Figures 9 and 10. First, as is outlined in more detail below, a single target probe is hybridized with a target nucleic acid. Each terminus of the probe hybridizes adjacently on the target nucleic acid and the OLA assay as described above occurs. When ligated, the probe is circularized while hybridized to the target nucleic acid. Addition of a polymerase results in extension of the circular probe. However, since the probe has no terminus, the polymerase continues to extend the probe repeatedly. Thus results in amplification of the circular probe.

A second alternative approach involves a two step process. In this embodiment, two ligation probes are initially ligated together, each containing a universal priming sequence. A rolling circle primer is then added, which has portions that will hybridize to the universal priming sequences. The presence of the ligase then causes the original probe to circularize, using the rolling circle primer as the polymerase primer, which is then amplified as above.

These embodiments also have the advantage that unligated probes need not necessarily be removed, as in the absence of the target, no significant amplification will occur. These benefits may be maximized by the design of the probes; for example, in the first embodiment, when there is a single target probe, placing the universal priming site close to the 5' end of the probe since this will only serve to generate short, truncated pieces, without adapters, in the absence of the ligation reaction.

Accordingly, in an preferred embodiment, a single oligonucleotide is used both for OLA and as the circular template for RCA (referred to herein as a "padlock probe" or a "RCA probe"). That is, each terminus of the oligonucleotide contains sequence complementary to the target nucleic acid and functions as an OLA primer as described above. That is, the first end of the RCA probe is substantially complementary to a first target domain, and the second end of the RCA probe is substantially complementary to a second target domain, adjacent to the first domain. Hybridization of the oligonucleotide to the target nucleic acid results in the formation of a hybridization complex. Ligation of the "primers" (which are the discrete ends of a single oligonucleotide) results in the formation of a modified hybridization complex containing a circular probe i.e. an RCA template complex. That is, the oligonucleotide is circularized while

still hybridized with the target nucleic acid. This serves as a circular template for RCA. Addition of a primer and a polymerase to the RCA template complex results in the formation of an amplicon.

Labeling of the amplicon can be accomplished in a variety of ways; for example, the polymerase may incorporate labeled nucleotides, or alternatively, a label probe is used that is substantially complementary to a portion of the RCA probe and comprises at least one label is used, as is generally outlined herein.

The polymerase can be any polymerase, but is preferably one lacking 3' exonuclease activity (3' exo$^-$). Examples of suitable polymerase include but are not limited to exonuclease minus DNA Polymerase I large (Klenow) Fragment, Phi29 DNA polymerase, Taq DNA Polymerase and the like. In addition, in some embodiments, a polymerase that will replicate single-stranded DNA (i.e. without a primer forming a double stranded section) can be used. In addition, while some embodiments utilize ligase, such as in the OLA or RCA, in some embodiments amplification alone is preferred. That is amplification is performed without a ligase step and without including a ligase enzyme.

In a preferred embodiment, the RCA probe contains an adapter sequence as outlined herein, with adapter capture probes on the array, for example on a microsphere when microsphere arrays are being used. Alternatively, unique portions of the RCA probes, for example all or part of the sequence corresponding to the target sequence, can be used to bind to a capture probe.

In a preferred embodiment, the padlock probe contains a restriction site. The restriction endonuclease site allows for cleavage of the long concatamers that are typically the result of RCA into smaller individual units that hybridize either more efficiently or faster to surface bound capture probes. Thus, following RCA, the product nucleic acid is contacted with the appropriate restriction endonuclease. This results in cleavage of the product nucleic acid into smaller fragments. The fragments are then hybridized with the capture probe that is immobilized resulting in a concentration of product fragments onto the microsphere. Again, as outlined herein, these fragments can be detected in one of two ways: either labelled nucleotides are incorporated during the replication step, or an additional label probe is added.

Thus, in a preferred embodiment, the padlock probe comprises a label sequence; i.e. a sequence that can be used to bind label probes and is substantially complementary to a label probe. In one embodiment, it is possible to use the same label sequence and label probe for all padlock probes on an array; alternatively, each padlock probe can have a different label sequence.

The padlock probe also contains a priming site for priming the RCA reaction. That is, each padlock probe comprises a sequence to which a primer nucleic acid hybridizes forming a template for the polymerase. The primer can be found in any portion of the circular probe. In a preferred embodiment, the primer is located at a discrete site in the probe. In this embodiment, the primer site in each distinct padlock probe is identical, e.g. is a universal priming site, although this is not required. Advantages of using primer sites with identical sequences include the ability to use only a single primer oligonucleotide to prime the RCA assay with a plurality of different hybridization complexes. That is, the padlock probe hybridizes uniquely to the target nucleic acid to which it is designed. A single primer hybridizes to all of the unique hybridization complexes forming a priming site for the polymerase. RCA then proceeds from an identical locus within each unique padlock probe of the hybridization complexes.

In an alternative embodiment, the primer site can overlap, encompass, or reside within any of the above-described elements of the padlock probe. That is, the primer can be found, for example, overlapping or within the restriction site or the identifier sequence. In this embodiment, it is necessary that the primer nucleic acid is designed to base pair with the chosen primer site.

Thus, the padlock probe of the invention contains at each terminus, sequences corresponding to OLA primers. The intervening sequence of the padlock probe contain in no particular order, an adapter sequence and a restriction endonuclease site. In addition, the padlock probe contains a RCA priming site.

Thus, in a preferred embodiment the OLA/RCA is performed in solution followed by restriction endonuclease cleavage of the RCA product. The cleaved product is then applied to an array comprising beads, each bead comprising a probe complementary to the adapter sequence located in the padlock probe. The amplified adapter sequence correlates with a particular target nucleic acid. Thus the incorporation of an endonuclease site allows the generation of short, easily hybridizable sequences. Furthermore, the unique adapter sequence in each rolling circle padlock probe sequence allows diverse sets of nucleic acid sequences to be analyzed in parallel on an array, since each sequence is resolved on the basis of hybridization specificity.

Thus, the present invention provides for the generation of amplicons (sometimes referred to herein as secondary targets).

In a preferred embodiment, the amplicons are labeled with a detection label. By "detection label" or "detectable label" herein is meant a moiety that allows detection. This may be a

primary label or a secondary label. Accordingly, detection labels may be primary labels (i.e. directly detectable) or secondary labels (indirectly detectable).

In a preferred embodiment, the detection label is a primary label. A primary label is one that can be directly detected, such as a fluorophore. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) magnetic, electrical, thermal labels; and c) colored or luminescent dyes. Labels can also include enzymes (horseradish peroxidase, etc.) and magnetic particles. Preferred labels include chromophores or phosphors but are preferably fluorescent dyes. Suitable dyes for use in the invention include, but are not limited to, fluorescent lanthanide complexes, including those of Europium and Terbium, fluorescein, rhodamine, tetramethylrhodamine, eosin, erythrosin, coumarin, methyl-coumarins, quantum dots (also referred to as "nanocrystals": see U.S.S.N. 09/315,584, hereby incorporated by reference), pyrene, Malacite green, stilbene, Lucifer Yellow, Cascade Blue™, Texas Red, Cy dyes (Cy3, Cy5, etc.), alexa dyes, phycoerythin, bodipy, and others described in the 6th Edition of the Molecular Probes Handbook by Richard P. Haugland, hereby expressly incorporated by reference.

In a preferred embodiment, a secondary detectable label is used. A secondary label is one that is indirectly detected; for example, a secondary label can bind or react with a primary label for detection, can act on an additional product to generate a primary label (e.g. enzymes), or may allow the separation of the compound comprising the secondary label from unlabeled materials, etc. Secondary labels include, but are not limited to, one of a binding partner pair such as biotin/streptavidin; chemically modifiable moieties; nuclease inhibitors, enzymes such as horseradish peroxidase, alkaline phosphatases, lucifierases, etc.

In a preferred embodiment, the secondary label is a binding partner pair. For example, the label may be a hapten or antigen, which will bind its binding partner. In a preferred embodiment, the binding partner can be attached to a solid support to allow separation of extended and non-extended primers. For example, suitable binding partner pairs include, but are not limited to: antigens (such as proteins (including peptides)) and antibodies (including fragments thereof (FAbs, etc.)); proteins and small molecules, including biotin/streptavidin; enzymes and substrates or inhibitors; other protein-protein interacting pairs; receptor-ligands; and carbohydrates and their binding partners. Nucleic acid - nucleic acid binding proteins pairs are also useful. In general, the smaller of the pair is attached to the NTP for incorporation into the primer. Preferred binding partner pairs include, but are not limited to, biotin (or imino-biotin) and streptavidin, digeoxinin and Abs, and Prolinx™ reagents (see www.prolinxinc.com/ie4/home.hmtl).

In a preferred embodiment, the binding partner pair comprises biotin or imino-biotin and

streptavidin. Imino-biotin is particularly preferred as imino-biotin disassociates from streptavidin in pH 4.0 buffer while biotin requires harsh denaturants (e.g. 6 M guanidinium HCl, pH 1.5 or 90% formamide at 95°C).

In a preferred embodiment, the binding partner pair comprises a primary detection label (for example, attached to the NTP and therefore to the amplicon) and an antibody that will specifically bind to the primary detection label. By "specifically bind" herein is meant that the partners bind with specificity sufficient to differentiate between the pair and other components or contaminants of the system. The binding should be sufficient to remain bound under the conditions of the assay, including wash steps to remove non-specific binding. In some embodiments, the dissociation constants of the pair will be less than about 10^{-4}-10^{-6} M^{-1}, with less than about 10^{-5} to 10^{-9} M^{-1} being preferred and less than about 10^{-7} -10^{-9} M^{-1} being particularly preferred.

In a preferred embodiment, the secondary label is a chemically modifiable moiety. In this embodiment, labels comprising reactive functional groups are incorporated into the nucleic acid. The functional group can then be subsequently labeled with a primary label. Suitable functional groups include, but are not limited to, amino groups, carboxy groups, maleimide groups, oxo groups and thiol groups, with amino groups and thiol groups being particularly preferred. For example, primary labels containing amino groups can be attached to secondary labels comprising amino groups, for example using linkers as are known in the art; for example, homo-or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference).

As outlined herein, labeling can occur in a variety of ways, as will be appreciated by those in the art. In general, labeling can occur in one of three ways: labels are incorporated into primers such that the amplification reaction results in amplicons that comprise the labels; labels are attached to dNTPs and incorporated by the polymerase into the amplicons; or the amplicons comprise a label sequence that is used to hybridize a label probe, and the label probe comprises the labels. It should be noted that in the latter case, the label probe can be added either before the amplicons are contacted with an array or afterwards.

A preferred embodiment utilizes one primer comprising a biotin, that is used to bind a fluorescently labeled streptavidin.

In a preferred embodiment following amplification, the amplicons are subjected to an additional amplification step. Preferably the additional amplification step is a T7 RNA polymerase reaction, although T7 amplification also can be the primary amplification step.

The advantage of following the amplification step with an additional amplification step such as the T7 RNA Polymerase reaction is that up to one hundred fold or more nucleic acid is generated therefore increasing the level of multiplexing.
As described above, the probes include T7 RNA polymerase priming sites for this additional step. Following amplification with T7 RNA polymerase, the resulting RNA contains a zip code and a universal primer that is allele specific. The resulting material is then detected.

In one embodiment the amplicons are detected by hybridization to an array. The array can be an ordered array or a random array as described herein. In addition, the array can be a liquid array. That is, the array can be a solution-phase array and detection is accomplished in a FACS, for example. In a preferred embodiment the detection array is a random BeadArray™.

In addition to the methods outlined herein, the present invention also provides methods for accomplishing genotyping of nucleic acids, including cDNA and genomic DNA. In general, this method can be described as follows, as is generally described in WO 00/63437, hereby expressly incorporated by reference. Genomic DNA is prepared from sample cells (and generally cut into smaller segments, for example through shearing or enzymatic treatment with enzymes such as DNAse I, as is well known in the art). Using any number of techniques, as are outlined below, the genomic fragments are attached, either covalently or securely, to a support such as beads or reaction wells (eppendorf tubes, microtiter wells, etc.). Any number of different genotyping reactions can then be done as outlined below, and the reaction products from these genotyping reactions are released from the support, amplified as necessary and added to an array of capture probes as outlined herein. In general, the methods described herein relate to the detection of nucleotide substitutions, although as will be appreciated by those in the art, deletions, insertions, inversions, etc. may also be detected. Universal primers can also be included as necessary.

These genotyping techniques fall into five general categories: (1) techniques that rely on traditional hybridization methods that utilize the variation of stringency conditions (temperature, buffer conditions, etc.) to distinguish nucleotides at the detection position; (2) extension techniques that add a base ("the base") to basepair with the nucleotide at the detection position; (3) ligation techniques, that rely on the specificity of ligase enzymes (or, in some cases, on the specificity of chemical techniques), such that ligation reactions occur preferentially if perfect complementarity exists at the detection position; (4) cleavage techniques, that also rely on enzymatic or chemical specificity such that cleavage occurs preferentially if perfect complementarity exists; and (5) techniques that combine these methods. See generally WO 00/63437, incorporated by reference in its entirety.

As above, if required, the target genomic sequence is prepared using known techniques, and

then attached to a solid support as defined herein. These techniques include, but are not limited to, enzymatic attachment, chemical attachment, photochemistry or thermal attachment and absorption.

In a preferred embodiment, as outlined herein, enzymatic techniques are used to attach the genomic DNA to the support. For example, terminal transferase end-labeling techniques can be used as outlined above; see Hermanson, Bioconjugate Techniques, San Diego, Academic Press, pp 640-643). In this embodiment, a nucleotide labeled with a secondary label (e.g. a binding ligand) is added to a terminus of the genomic DNA; supports coated or containing the binding partner can thus be used to immobilize the genomic DNA. Alternatively, the terminal transferase can be used to add nucleotides with special chemical functionalities that can be specifically coupled to a support. Similarly, random-primed labeling or nick-translation labeling (supra, pp. 640-643) can also be used.

In a preferred embodiment, chemical labeling (supra, pp.6444-671) can be used. In this embodiment, bisulfite-catalyzed transamination, sulfonation of cytosine residues, bromine activation of T, C and G bases, periodate oxidation of RNA or carbodiimide activation of 5' phosphates can be done.

In a preferred embodiment, photochemistry or heat-activated labeling is done (supra, p162-166). Thus for example, aryl azides and nitrenes preferably label adenosines, and to a less extent C and T (Aslam et al., Bioconjugation: Protein Coupling Techniques for Biomedical Sciences; New York, Grove's Dictionaries, 833 pp.). Psoralen or angelicin compounds can also be used (Aslam, p492, supra). The preferential modification of guanine can be accomplished via intercalation of platinum complexes (Aslam, supra).

In a preferred embodiment, the genomic DNA can be absorbed on positively charged surfaces, such as an amine coated solid phase. The genomic DNA can be cross-linked to the surface after physical absorption for increased retention (e.g. PEI coating and glutaraldehyde cross-linking; Aslam, supra, p.485).

In a preferred embodiment, direct chemical attached or photocrosslinking can be done to attach the genomic DNA to the solid phase, by using direct chemical groups on the solid phase substrate. For example, carbodiimide activation of 5' phosphates, attachment to exocyclic amines on DNA bases, and psoralen can be attached to the solid phase for crosslinking to the DNA.

Once added to the support, the target genomic sequence can be used in a variety of reactions for a variety of reasons. For example, in a preferred embodiment, genotyping reactions are

done. Similarly, these reactions can also be used to detect the presence or absence of a target genomic sequence. In addition, in any reaction, quantitation of the amount of a target genomic sequence may be done. While the discussion below focuses on genotyping reactions, the discussion applies equally to detecting the presence of target sequences and/or their quantification.

As will be appreciated by those in the art, the reactions described below can take on a wide variety of formats. In one embodiment, genomic DNA is attached to a solid support, and probes comprising universal primers are added to form hybridization complexes, in a variety of formats as outlined herein. The non-hybridized probes are then removed, and the hybridization complexes are denatured This releases the probes (which frequently have been altered in some way). They are then amplified and added to an array of capture probes. In a preferred embodiment, non-hybridized primers are removed prior to the enzymatic step. Several embodiments of this have been described above. Alternatively, genomic DNA is attached to a solid support, and genotyping reactions are done in formats that can allow amplification as well, either during the genotyping reaction (e.g. through the use of heat cycling) or after, without the use of universal primers. Thus, for example, when labeled probes are used, they can be hybridized to the immobilized genomic DNA, unbound materials removed, and then eluted and collected to be added to arrays. This may be repeated for amplification purposes, with the elution fractions pooled and added to the array. In addition, alternative amplification schemes such as extending a product of the invasive cleavage reaction (described below) to include universal primers or universal primers and adapters can be performed. In one embodiment this allows the reuse of immobilized target sequences with a different set or sets of target probes.

In some embodiments, amplification of the product of the genotyping reactions is not necessary. For example, in genomes of less complexity, e.g. bacterial , yeast and *Drosophila*, detectable signal is achieved without the need for amplification. This is particularly true when primer extension is performed and more than one base is added to the probe, as is more fully outlined below.

In a preferred embodiment, straight hybridization methods are used to elucidate the identity of the base at the detection position. Generally speaking, these techniques break down into two basic types of reactions: those that rely on competitive hybridization techniques, and those that discriminate using stringency parameters and combinations thereof.

In a preferred embodiment, the use of competitive hybridization probes is done to elucidate either the identity of the nucleotide(s) at the detection position or the presence of a mismatch. For example, sequencing by hybridization has been described (Drmanac et al., Genomics

4:114 (1989); Koster et al., Nature Biotechnology 14:1123 (1996); U.S. Patent Nos. 5,525,464; 5,202,231 and 5,695,940, among others, all of which are hereby expressly incorporated by reference in their entirety).

As outlined above, in a preferred embodiment, a plurality of readout probes are used to identify the base at the detection position. In this embodiment, each different readout probe comprises either a different detection label (which, as outlined below, can be either a primary label or a secondary label) or a different adapter, and a different base at the position that will hybridize to the detection position of the target sequence (herein referred to as the readout position) such that differential hybridization will occur.

Accordingly, in some embodiments, a detectable label is incorporated into the readout probe. In a preferred embodiment, a set of readout probes are used, each comprising a different base at the readout position. In some embodiments, each readout probe comprises a different label, that is distinguishable from the others. For example, a first label may be used for probes comprising adenosine at the readout position, a second label may be used for probes comprising guanine at the readout position, etc. In a preferred embodiment, the length and sequence of each readout probe is identical except for the readout position, although this need not be true in all embodiments.

The number of readout probes used will vary depending on the end use of the assay. For example, many SNPs are biallelic, and thus two readout probes, each comprising an interrogation base that will basepair with one of the detection position bases. For sequencing, for example, for the discovery of SNPs, a set of four readout probes are used, although SNPs may also be discovered with fewer readout parameters.

In one embodiment, the probes used as readout probes are "Molecular Beacon" probes as are generally described in Whitcombe et al., Nature Biotechnology 17:804 (1999), hereby incorporated by reference. As is known in the art, Molecular Beacon probes form "hairpin" type structures, with a fluorescent label on one end and a quencher on the other. In the absence of the target sequence, the ends of the hairpin hybridize, causing quenching of the label. In the presence of a target sequence, the hairpin structure is lost in favor of target sequence binding, resulting in a loss of quenching and thus an increase in signal.

In a preferred embodiment, extension genotyping is done. In this embodiment, any number of techniques are used to add a nucleotide to the readout position of a probe hybridized to the target sequence adjacent to the detection position. By relying on enzymatic specificity, preferentially a perfectly complementary base is added. All of these methods rely on the enzymatic incorporation of nucleotides at the detection position. This may be done using

chain terminating dNTPs, such that only a single base is incorporated (e.g. single base extension methods), or under conditions that only a single type of nucleotide is added followed by identification of the added nucleotide (extension and pyrosequencing techniques).

In a preferred embodiment, single base extension (SBE; sometimes referred to as "minisequencing") is used to determine the identity of the base at the detection position. SBE utilizes an extension primer with at least one adapter sequence that hybridizes to the target nucleic acid immediately adjacent to the detection position, to form a hybridization complex. A polymerase (generally a DNA polymerase) is used to extend the 3' end of the primer with a nucleotide analog labeled with a detection label as described herein. Based on the fidelity of the enzyme, a nucleotide is only incorporated into the readout position of the growing nucleic acid strand if it is perfectly complementary to the base in the target strand at the detection position. The nucleotide may be derivatized such that no further extensions can occur, so only a single nucleotide is added. Once the labeled nucleotide is added, detection of the label proceeds as outlined herein. Again, amplification in this case is accomplished through cycling or repeated rounds of reaction/elution, although in some embodiments amplification is not necessary.

The reaction is initiated by introducing the hybridization complex comprising the target genomic sequence on the support to a solution comprising a first nucleotide. In general, the nucleotides comprise a detectable label, which may be either a primary or a secondary label. In addition, the nucleotides may be nucleotide analogs, depending on the configuration of the system. For example, if the dNTPs are added in sequential reactions, such that only a single type of dNTP can be added, the nucleotides need not be chain terminating. In addition, in this embodiment, the dNTPs may all comprise the same type of label.

Alternatively, if the reaction comprises more than one dNTP, the dNTPs should be chain terminating, that is, they have a blocking or protecting group at the 3' position such that no further dNTPs may be added by the enzyme. As will be appreciated by those in the art, any number of nucleotide analogs may be used, as long as a polymerase enzyme will still incorporate the nucleotide at the readout position. Preferred embodiments utilize dideoxy-triphosphate nucleotides (ddNTPs) and halogenated dNTPs. Generally, a set of nucleotides comprising ddATP, ddCTP, ddGTP and ddTTP is used, each with a different detectable label, although as outlined herein, this may not be required. Alternative preferred embodiments use acyclo nucleotides (NEN). These chain terminating nucleotide analogs are particularly good substrates for Deep vent (exo$^-$) and thermosequenase.

In addition, as will be appreciated by those in the art, the single base extension reactions of the present invention allow the precise incorporation of modified bases into a growing nucleic

acid strand. Thus, any number of modified nucleotides may be incorporated for any number of reasons, including probing structure-function relationships (e.g. DNA:DNA or DNA:protein interactions), cleaving the nucleic acid, crosslinking the nucleic acid, incorporate mismatches, etc.

As will be appreciated by those in the art, the configuration of the genotyping SBE system can take on several forms.

In addition, since unextended primers do not comprise labels, the unextended primers need not be removed. However, they may be, if desired, as outlined below; for example, if a large excess of primers are used, there may not be sufficient signal from the extended primers competing for binding to the surface.

Alternatively, one of skill in the art could use a single label and temperature to determine the identity of the base; that is, the readout position of the extension primer hybridizes to a position on the capture probe. However, since the three mismatches will have lower Tms than the perfect match, the use of temperature could elucidate the identity of the detection position base.

In a preferred embodiment, the determination of the identity of the base at the detection position of the target sequence proceeds using invasive cleavage technology. As outlined above for amplification, invasive cleavage techniques rely on the use of structure-specific nucleases, where the structure can be formed as a result of the presence or absence of a mismatch. Generally, invasive cleavage technology may be described as follows. A target nucleic acid is recognized by two distinct probes. A first probe, generally referred to herein as an "invader" probe, is substantially complementary to a first portion of the target nucleic acid. A second probe, generally referred to herein as a "signal probe", is partially complementary to the target nucleic acid; the 3' end of the signal oligonucleotide is substantially complementary to the target sequence while the 5' end is non-complementary and preferably forms a single-stranded "tail" or "arm". The non-complementary end of the second probe preferably comprises a "generic" or "unique" sequence, frequently referred to herein as a "detection sequence", that is used to indicate the presence or absence of the target nucleic acid, as described below. The detection sequence of the second probe may comprise at least one detectable label (for cycling purposes), or preferably comprises one or more universal priming sites and/or an adapter sequence. Alternative methods have the detection sequence functioning as a target sequence for a capture probe, and thus rely on sandwich configurations using label probes.

Hybridization of the first and second oligonucleotides near or adjacent to one another on the

target genomic nucleic acid forms a number of structures.

Accordingly, the present invention provides methods of determining the identity of a base at the detection position of a target sequence. In this embodiment, the target sequence comprises, 5' to 3', a first target domain comprising an overlap domain comprising at least a nucleotide in the detection position, and a second target domain contiguous with the detection position. A first probe (the "invader probe") is hybridized to the first target domain of the target sequence. A second probe (the "signal probe"), comprising a first portion that hybridizes to the second target domain of the target sequence and a second portion that does not hybridize to the target sequence, is hybridized to the second target domain. If the second probe comprises a base that is perfectly complementary to the detection position a cleavage structure is formed. The addition of a cleavage enzyme, such as is described in U.S. Patent Nos. 5,846,717; 5,614,402; 5,719,029; 5,541,311 and 5,843,669, all of which are expressly incorporated by reference, results in the cleavage of the detection sequence from the signalling probe. This then can be used as a target sequence in an assay complex.

In addition, as for a variety of the techniques outlined herein, unreacted probes (i.e. signalling probes, in the case of invasive cleavage), may be removed using any number of techniques. For example, the use of a binding partner coupled to a solid support comprising the other member of the binding pair can be done. Similarly, after cleavage of the primary signal probe, the newly created cleavage products can be selectively labeled at the 3' or 5' ends using enzymatic or chemical methods.

Again, as outlined above, the detection of the invasive cleavage reaction can occur directly, in the case where the detection sequence comprises at least one label, or indirectly, using sandwich assays, through the use of additional probes; that is, the detection sequences can serve as target sequences, and detection may utilize amplification probes, capture probes, capture extender probes, label probes, and label extender probes, etc. In one embodiment, a second invasive cleavage reaction is performed on solid-phase thereby making it easier perform multiple reactions.

In addition, as for most of the techniques outlined herein, these techniques may be done for the two strands of a double-stranded target sequence. The target sequence is denatured, and two sets of probes are added: one set as outlined above for one strand of the target, and a separate set for the other strand of the target.

Thus, the invasive cleavage reaction requires, in no particular order, an invader probe, a signalling probe, and a cleavage enzyme.

It is also possible to combine two or more of these techniques to do genotyping, quantification, detection of sequences, etc., again as outlined in WO 00/63437, expressly incorporated by reference, including combinations of competitive hybridization and extension, particularly SBE; a combination of competitive hybridization and invasive cleavage; invasive cleavage and ligation; a combination of invasive cleavage and extension reactions; a combination of OLA and SBE; a combination of OLA and PCR; a combination of competitive hybridization and ligation; and a combination of competitive hybridization and invasive cleavage.

Solution phase assays

Alternatively, the assays of the invention can be run in solution, followed by detection of the amplicons, either by the addition of the amplicons to an array or utilizing other methods as outlined herein (mass spectroscopy, electrophoresis, etc.) as outlined herein. In this embodiment, a variety of methods can be used to remove unhybridized target probes, as outlined in WO 00/63437, expressly incorporated by reference herein.

For example, if the target analyte is not immobilized, separation methods based on the differences between single-stranded and double-stranded nucleic acids may be done. For example, there are a variety of double-stranded specific moieties known, that preferentially interact with double-stranded nucleic acids over single stranded nucleic acids. For example, there are a wide variety of intercalators known, that insert into the stacked basepairs of double stranded nucleic acid. Two of the best known examples are ethidium bromide and actinomycin D. Similarly, there are a number of major groove and minor groove binding proteins which can be used to distinguish between single stranded and double stranded nucleic acids. Similar to the poly(T) embodiment, these moieties can be attached to a support such as magnetic beads and used to preferentially bind the hybridization complexes, to remove the non-hybridized target probes and target sequences during washing steps. The hybridization complexes are then released from the beads using a denaturation step such as a thermal step.

In the case where the OLA reaction is done, an additional embodiment, depicted in Figure 8, may be done to remove unhybridized primers. In this embodiment, a nuclease inhibitor is added to the 3' end of the downstream ligation probe, which does not comprise the adapter sequence. Thus, any nucleic acids that do not contain the inhibitors (including both the 5' unligated probe and the target sequences themselves) will be digested upon addition of a 3'-exonuclease. The ligation products are protected from exo I digestion by including, for example, 4-phosphorothioate residues at their 3' terminus, thereby, rendering them resistant to exonuclease digestion. The unligated detection oligonucleotides are not protected and are digested. Since the 5' upstream ligation probe carries the adapter sequence, the unligated downstream probe, which does carry the nuclease inhibitor and is thus also not digested, does

not bind to the array and can be washed away. The nuclease inhibitors may also be used in non-OLA utilities as well.

Suitable nuclease inhibitors are known in the art and comprise thiol nucleotides. In this embodiment, suitable 3'-exonucleases include, but are not limited to, exo I, exo III, exo VII, and 3'-5' exophosphodiesterases.

Following the amplification procedure, there is present sufficient nucleic acid material to detect the results of the genotyping assays through conventional means. In the preferred embodiment, the probes used in the mutation detection reaction also contain address sequences. During the amplification process, the address sequences used to read out the results are simultaneously amplified with the mutation-detection probes. When the amplified material is applied to a detection substrate, such as an array where complementary address sequences are provided, the amplified nucleic acid probes are then detected by known methods.

In the preferred method, the detection substrate is a random array substrate, as described in U.S. Patent No.6,023,540 which is incorporated by reference herein, where the hybridization of complementary nucleic acid sequences, or address sequences, are used as the particular detection means. The arrays can be manufactured with a standard set of nucleic acid address sequences, one address sequence for each different nucleic acid to be detected. The complementary nucleic acid sequences are provided as part of the linear nucleic acid sequences of the mutation-detection probes, inside of the working portion of the amplification primers. During amplification, the address sequences are amplified along with each respective mutation-detection probe. In order to detect the results of the multiplexed genotyping reaction, the resulting amplified mutation-detection probe mixture is applied to the array, whereby the complementary address sequences on the mutation-detection probes and on the array hybridize, and the results are analyzed by known methods, such as fluorescence.

Other detection schemes such as flow cytometry, mass spectroscopy, spotted arrays, or spatially-directed arrays can also be used to simultaneously read the results of the multiplexed nucleic acid detection reactions.

Accordingly, the present invention provides methods and compositions useful in the detection of nucleic acids, particularly the labeled amplicons outlined herein. As is more fully outlined below, preferred systems of the invention work as follows. Amplicons are attached (via hybridization) to an array site. This attachment can be either directly to a capture probe on the surface, through the use of adapters, or indirectly, using capture extender probes as outlined herein. In some embodiments, the target sequence itself comprises the labels. Alternatively,

a label probe is then added, forming an assay complex. The attachment of the label probe may be direct (i.e. hybridization to a portion of the target sequence), or indirect (i.e. hybridization to an amplifier probe that hybridizes to the target sequence), with all the required nucleic acids forming an assay complex.

Accordingly, the present invention provides array compositions comprising at least a first substrate with a surface comprising individual sites. By "array" or "biochip" herein is meant a plurality of nucleic acids in an array format; the size of the array will depend on the composition and end use of the array. Nucleic acids arrays are known in the art, and can be classified in a number of ways; both ordered arrays (e.g. the ability to resolve chemistries at discrete sites), and random arrays are included. Ordered arrays include, but are not limited to, those made using photolithography techniques (Affymetrix GeneChip™), spotting techniques (Synteni and others), printing techniques (Hewlett Packard and Rosetta), three dimensional "gel pad" arrays, etc. A preferred embodiment utilizes microspheres on a variety of substrates including fiber optic bundles, as are outlined in PCTs US98/21193, PCT US99/14387 and PCT US98/05025; WO98/50782; and U.S.S.N.s 09/287,573, 09/151,877, 09/256,943, 09/316,154, 60/119,323, 09/315,584; all of which are expressly incorporated by reference.

Arrays containing from about 2 different bioactive agents (e.g. different beads, when beads are used) to many millions can be made, with very large arrays being possible. Generally, the array will comprise from two to as many as a billion or more, depending on the size of the beads and the substrate, as well as the end use of the array, thus very high density, high density, moderate density, low density and very low density arrays may be made. Preferred ranges for very high density arrays are from about 10,000,000 to about 2,000,000,000, with from about 100,000,000 to about 1,000,000,000 being preferred (all numbers being in square cm). High density arrays range about 100,000 to about 10,000,000, with from about 1,000,000 to about 5,000,000 being particularly preferred. Moderate density arrays range from about 10,000 to about 100,000 being particularly preferred, and from about 20,000 to about 50,000 being especially preferred. Low density arrays are generally less than 10,000, with from about 1,000 to about 5,000 being preferred. Very low density arrays are less than 1,000, with from about 10 to about 1000 being preferred, and from about 100 to about 500 being particularly preferred. In some embodiments, the compositions of the invention may not be in array format; that is, for some embodiments, compositions comprising a single bioactive agent may be made as well. In addition, in some arrays, multiple substrates may be used, either of different or identical compositions. Thus for example, large arrays may comprise a plurality of smaller substrates.

In addition, one advantage of the present compositions is that particularly through the use of

fiber optic technology, extremely high density arrays can be made. Thus for example, because beads of 200 μm or less (with beads of 200 nm possible) can be used, and very small fibers are known, it is possible to have as many as 40,000 or more (in some instances, 1 million) different elements (e.g. fibers and beads) in a 1 mm^2 fiber optic bundle, with densities of greater than 25,000,000 individual beads and fibers (again, in some instances as many as 50-100 million) per 0.5 cm^2 obtainable (4 million per square cm for 5 μ center-to-center and 100 million per square cm for 1 μ center-to-center).

By "substrate", "array substrate" or "solid support" or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of beads and is amenable to at least one detection method. It should be noted that the array substrate is distinct from the "capture surface" described above. The capture surface is for the immobilization of target nucleic acids while the array substrate is for detection of amplicons, i.e. the results of the detection or genotyping assay. As will be appreciated by those in the art, the number of possible array substrates is very large. Possible array substrates include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, optical fiber bundles, and a variety of other polymers. In general, the array substrates allow optical detection and do not themselves appreciably fluoresce.

Generally the array substrate is flat (planar), although as will be appreciated by those in the art, other configurations of substrates may be used as well; for example, three dimensional configurations can be used, for example by embedding the beads in a porous block of plastic that allows sample access to the beads and using a confocal microscope for detection. Similarly, the beads may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Preferred substrates include optical fiber bundles as discussed below, and flat planar substrates such as paper, glass, polystyrene and other plastics and acrylics.

In a preferred embodiment, the substrate is an optical fiber bundle or array, as is generally described in U.S.S.N.s 08/944,850 and 08/519,062, PCT US98/05025, and PCT US98/09163, all of which are expressly incorporated herein by reference. . Preferred embodiments utilize preformed unitary fiber optic arrays. By "preformed unitary fiber optic array" herein is meant an array of discrete individual fiber optic strands that are co-axially disposed and joined along their lengths. The fiber strands are generally individually clad. However, one thing that distinguished a preformed unitary array from other fiber optic formats is that the fibers are not

individually physically manipulatable; that is, one strand generally cannot be physically separated at any point along its length from another fiber strand.

Generally, the array of array compositions of the invention can be configured in several ways; see for example U.S.S.N. 09/473,904, hereby expressly incorporated by reference. In a preferred embodiment, as is more fully outlined below, a "one component" system is used. That is, a first substrate comprising a plurality of assay locations (sometimes also referred to herein as "assay wells"), such as a microtiter plate, is configured such that each assay location contains an individual array. That is, the assay location and the array location are the same. For example, the plastic material of the microtiter plate can be formed to contain a plurality of "bead wells" in the bottom of each of the assay wells. Beads containing the capture probes of the invention can then be loaded into the bead wells in each assay location as is more fully described below. Arrays are described in U.S. Patent No. 6,023,540 and U.S.S.N.'s 09/151,877, filed September 11, 1998, 09/450,829, filed November 29, 1999, 09/816,651, filed March 23, 2001, and 09/840,012, filed April 20, 2001, all of which are expressly incorporated herein by reference. In addition, other arrays are described in 60/181,631, filed February 10, 2000, 09/782,588, filed February 12, 2001, 60/113,968, filed December 28, 1998, 090/256,943, filed February 24, 1999, 09/473,904, filed December 28, 1999 and 09/606,369, filed June 28, 2000, all of which are expressly incorporated herein by reference.

Alternatively, a "two component" system can be used. In this embodiment, the individual arrays are formed on a second substrate, which then can be fitted or "dipped" into the first microtiter plate substrate. A preferred embodiment utilizes fiber optic bundles as the individual arrays, generally with "bead wells" etched into one surface of each individual fiber, such that the beads containing the capture probes are loaded onto the end of the fiber optic bundle. The composite array thus comprises a number of individual arrays that are configured to fit within the wells of a microtiter plate.

By "composite array" or "combination array" or grammatical equivalents herein is meant a plurality of individual arrays, as outlined above. Generally the number of individual arrays is set by the size of the microtiter plate used; thus, 96 well, 384 well and 1536 well microtiter plates utilize composite arrays comprising 96, 384 and 1536 individual arrays, although as will be appreciated by those in the art, not each microtiter well need contain an individual array. It should be noted that the composite arrays can comprise individual arrays that are identical, similar or different. That is, in some embodiments, it may be desirable to do the same 2,000 assays on 96 different samples; alternatively, doing 192,000 experiments on the same sample (i.e. the same sample in each of the 96 wells) may be desirable. Alternatively, each row or column of the composite array could be the same, for redundancy/quality control. As will be appreciated by those in the art, there are a variety of ways to configure the system. In

addition, the random nature of the arrays may mean that the same population of beads may be added to two different surfaces, resulting in substantially similar but perhaps not identical arrays.

At least one surface of the substrate is modified to contain discrete, individual sites for later association of microspheres. These sites may comprise physically altered sites, i.e. physical configurations such as wells or small depressions in the substrate that can retain the beads, such that a microsphere can rest in the well, or the use of other forces (magnetic or compressive), or chemically altered or active sites, such as chemically functionalized sites, electrostatically altered sites, hydrophobically/ hydrophilically functionalized sites, spots of adhesive, etc.

The sites may be a pattern, i.e. a regular design or configuration, or randomly distributed. A preferred embodiment utilizes a regular pattern of sites such that the sites may be addressed in the X-Y coordinate plane. "Pattern" in this sense includes a repeating unit cell, preferably one that allows a high density of beads on the substrate. However, it should be noted that these sites may not be discrete sites. That is, it is possible to use a uniform surface of adhesive or chemical functionalities, for example, that allows the attachment of beads at any position. That is, the surface of the substrate is modified to allow attachment of the microspheres at individual sites, whether or not those sites are contiguous or non-contiguous with other sites. Thus, the surface of the substrate may be modified such that discrete sites are formed that can only have a single associated bead, or alternatively, the surface of the substrate is modified and beads may go down anywhere, but they end up at discrete sites. That is, while beads need not occupy each site on the array, no more than one bead occupies each site.

In a preferred embodiment, the surface of the substrate is modified to contain wells, i.e. depressions in the surface of the substrate. This may be done as is generally known in the art using a variety of techniques, including, but not limited to, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those in the art, the technique used will depend on the composition and shape of the substrate.

In a preferred embodiment, physical alterations are made in a surface of the substrate to produce the sites. In a preferred embodiment, the substrate is a fiber optic bundle and the surface of the substrate is a terminal end of the fiber bundle, as is generally described in 08/818,199 and 09/151,877, both of which are hereby expressly incorporated by reference. In this embodiment, wells are made in a terminal or distal end of a fiber optic bundle comprising individual fibers. In this embodiment, the cores of the individual fibers are etched, with respect to the cladding, such that small wells or depressions are formed at one end of the fibers. The

required depth of the wells will depend on the size of the beads to be added to the wells.

Generally in this embodiment, the microspheres are non-covalently associated in the wells, although the wells may additionally be chemically functionalized as is generally described below, cross-linking agents may be used, or a physical barrier may be used, i.e. a film or membrane over the beads.

In a preferred embodiment, the surface of the substrate is modified to contain chemically modified sites, that can be used to attach, either covalently or non-covalently, the microspheres of the invention to the discrete sites or locations on the substrate. "Chemically modified sites" in this context includes, but is not limited to, the addition of a pattern of chemical functional groups including amino groups, carboxy groups, oxo groups and thiol groups, that can be used to covalently attach microspheres, which generally also contain corresponding reactive functional groups; the addition of a pattern of adhesive that can be used to bind the microspheres (either by prior chemical functionalization for the addition of the adhesive or direct addition of the adhesive); the addition of a pattern of charged groups (similar to the chemical functionalities) for the electrostatic attachment of the microspheres, i.e. when the microspheres comprise charged groups opposite to the sites; the addition of a pattern of chemical functional groups that renders the sites differentially hydrophobic or hydrophilic, such that the addition of similarly hydrophobic or hydrophilic microspheres under suitable experimental conditions will result in association of the microspheres to the sites on the basis of hydroaffinity. For example, the use of hydrophobic sites with hydrophobic beads, in an aqueous system, drives the association of the beads preferentially onto the sites. As outlined above, "pattern" in this sense includes the use of a uniform treatment of the surface to allow attachment of the beads at discrete sites, as well as treatment of the surface resulting in discrete sites. As will be appreciated by those in the art, this may be accomplished in a variety of ways.

In some embodiments, the beads are not associated with a substrate. That is, the beads are in solution or are not distributed on a patterned substrate.

In a preferred embodiment, the compositions of the invention further comprise a population of microspheres. By "population" herein is meant a plurality of beads as outlined above for arrays. Within the population are separate subpopulations, which can be a single microsphere or multiple identical microspheres. That is, in some embodiments, as is more fully outlined below, the array may contain only a single bead for each capture probe; preferred embodiments utilize a plurality of beads of each type.

By "microspheres" or "beads" or "particles" or grammatical equivalents herein is meant small

discrete particles. The composition of the beads will vary, depending on the class of capture probe and the method of synthesis. Suitable bead compositions include those used in peptide, nucleic acid and organic moiety synthesis, including, but not limited to, plastics, ceramics, glass, polystyrene, methylstyrene, acrylic polymers, paramagnetic materials, thoria sol, carbon graphite, titanium dioxide, latex or cross-linked dextrans such as Sepharose, cellulose, nylon, cross-linked micelles and Teflon may all be used. *"Microsphere Detection Guide"* from Bangs Laboratories, Fishers IN is a helpful guide.

The beads need not be spherical; irregular particles may be used. In addition, the beads may be porous, thus increasing the surface area of the bead available for either capture probe attachment or tag attachment. The bead sizes range from nanometers, i.e. 100 nm, to millimeters, i.e. 1 mm, with beads from about 0.2 micron to about 200 microns being preferred, and from about 0.5 to about 5 micron being particularly preferred, although in some embodiments smaller beads may be used.

Each microsphere comprises a capture probe, although as will be appreciated by those in the art, there may be some microspheres which do not contain a capture probe, depending on the synthetic methods.

Attachment of the nucleic acids may be done in a variety of ways, as will be appreciated by those in the art, including, but not limited to, chemical or affinity capture (for example, including the incorporation of derivatized nucleotides such as AminoLink or biotinylated nucleotides that can then be used to attach the nucleic acid to a surface, as well as affinity capture by hybridization), cross-linking, and electrostatic attachment, etc. In a preferred embodiment, affinity capture is used to attach the nucleic acids to the beads. For example, nucleic acids can be derivatized, for example with one member of a binding pair, and the beads derivatized with the other member of a binding pair. Suitable binding pairs are as described herein for IBL/DBL pairs. For example, the nucleic acids may be biotinylated (for example using enzymatic incorporate of biotinylated nucleotides, for by photoactivated cross-linking of biotin). Biotinylated nucleic acids can then be captured on streptavidin-coated beads, as is known in the art. Similarly, other hapten-receptor combinations can be used, such as digoxigenin and anti-digoxigenin antibodies. Alternatively, chemical groups can be added in the form of derivatized nucleotides, that can them be used to add the nucleic acid to the surface.

Similarly, affinity capture utilizing hybridization can be used to attach nucleic acids to beads.

Alternatively, chemical crosslinking may be done, for example by photoactivated crosslinking of thymidine to reactive groups, as is known in the art.

In a preferred embodiment, each bead comprises a single type of capture probe, although a plurality of individual capture probes are preferably attached to each bead. Similarly, preferred embodiments utilize more than one microsphere containing a unique capture probe; that is, there is redundancy built into the system by the use of subpopulations of microspheres, each microsphere in the subpopulation containing the same capture probe.

As will be appreciated by those in the art, the capture probes may either be synthesized directly on the beads, or they may be made and then attached after synthesis. In a preferred embodiment, linkers are used to attach the capture probes to the beads, to allow both good attachment, sufficient flexibility to allow good interaction with the target molecule, and to avoid undesirable binding reactions.

In a preferred embodiment, the capture probes are synthesized directly on the beads. As is known in the art, many classes of chemical compounds are currently synthesized on solid supports, such as peptides, organic moieties, and nucleic acids. It is a relatively straightforward matter to adjust the current synthetic techniques to use beads.

In a preferred embodiment, the capture probes are synthesized first, and then covalently attached to the beads. As will be appreciated by those in the art, this will be done depending on the composition of the capture probes and the beads. The functionalization of solid support surfaces such as certain polymers with chemically reactive groups such as thiols, amines, carboxyls, etc. is generally known in the art. Accordingly, "blank" microspheres may be used that have surface chemistries that facilitate the attachment of the desired functionality by the user. Some examples of these surface chemistries for blank microspheres include, but are not limited to, amino groups including aliphatic and aromatic amines, carboxylic acids, aldehydes, amides, chloromethyl groups, hydrazide, hydroxyl groups, sulfonates and sulfates.

When random arrays or liquid arrays are used, an encoding/decoding system must be used. For example, when microsphere arrays are used, the beads are generally put onto the substrate randomly; as such there are several ways to correlate the functionality on the bead with its location, including the incorporation of unique optical signatures, generally fluorescent dyes, that could be used to identify the nucleic acid on any particular bead. This allows the synthesis of the capture probes to be divorced from their placement on an array, i.e. the capture probes may be synthesized on the beads, and then the beads are randomly distributed on a patterned surface. Since the beads are first coded with an optical signature, this means that the array can later be "decoded", i.e. after the array is made, a correlation of the location of an individual site on the array with the bead or probe at that particular site can be made. This means that the beads may be randomly distributed on the array, a fast and

inexpensive process as compared to either the in situ synthesis or spotting techniques of the prior art.

When liquid arrays are used, beads to which the amplicons are immobilized can be analyzed by FACS. Again, beads can be decoded to determine which amplicon is immobilized on the bead. This is an indication of the presence of the target analyte.

However, the drawback to these methods is that for a large array, the system requires a large number of different optical signatures, which may be difficult or time-consuming to utilize. Accordingly, methods for analysis and decoding of arrays are described in 08/944,850, filed October 6, 1997, PCT/US98/21193, filed October 6, 1998, 09/287,573, filed April 6, 1999, PCT/US00/09183, filed May 6, 2000, 60/238,866, filed October 6, 2000, 60/119,323, filed February 9, 1999, 09/500,555, filed February 9, 2000, 09/636,387, filed August 9, 2000, 60/151,483, filed August 30, 1999, 60/151,668, filed August 31, 1999, 09/651,181, filed August 30, 2000,60/272,803, filed March 1, 2001, all of which are expressly incorporated herein by reference. In addition, methods of decoding arrays are described in 60/090,473, filed June 24, 1998, 09/189,543, filed November 10, 1998, 09/344,526, filed June 24, 1999, PCT/US99/14387, filed June 24, 1999, 60/172, 106, filed December 23, 1999, 60/235,531, filed September 26, 2000, 09/748,706, filed December 22, 2000, and provisional application entitled Decoding of Array Sensors with Microspheres, filed June 28, 2001 (no serial number received), all of which are expressly incorporated herein by reference.

As outlined herein, the present invention finds use in a wide variety of applications. All references cited herein are incorporated by reference.

Examples

Attachment of genomic DNA to a solid support

1. Fragmentation of Genomic DNA

Human Genomic DNA	10 mg (100 µl)
10X DNase I Buffer	12.5 µl
DNase I (1 U/ µl, BRL)	0.5 µl
ddH2O	12 µl

Incubate 37°C for 10 min. Add 1.25 µl 0.5 M EDTA, Heat at 99°C for 15 min.

2. Precipitation of fragmented genomic DNA

DNase I fragmented genomic DNA	125 µl
Quick-Precip Plus Solution (Edge Biosystems)	20 µl
Cold 100% EtOH	300 µl

Store at –20°C for 20 min. Spin at 12,500 rpm for 5 min. Wash pellet 2x with 70% EtOH, and air dry.

3. Terminal Transferase End-Labeling with Biotin

DNase I fragmented and precipitated genomic DNA (in H2O)	77.3 µl
5X Terminal transferase buffer	20 µl
Biotin-N6-ddATP (1 mM, NEN)	1 µl
Terminal transferase (15 U/µl)	1.7 µl

37°C for 60 min. Add 1 µl 0.5 M EDTA, then heat at 99°C for 15 min

4. Precipitation of Biotin-labeled genomic DNA

Biotin-labeled genomic DNA	100 µl
Quick-Precip Solution	20 µl
EtOH	250 µl

-20°C for 20 min and spin at 12,500 rpm for 5 min, wash 2x with 70% EtOH and air dry.

5. Immobilization of Biotin-labeled Genomic DNA to Streptavidin-coated PCR tubes

Heat-denature genomic DNA for 10 min on 95°C heat block.

Biotin-labeled genomic DNA (0.3 µg/ µl)	3 µl
2 x binding buffer	25 µl
SNP Primers (50 nM)	10 µl
ddH2O	12 µl

Incubate at 60°C for 60 min.

Wash 1x with 1 X binding buffer,
1X with 1 X washing buffer,
1X with 1 X ligation buffer.

1X binding buffer: 20 mM Tris-HCl, pH7.5, 0.5M NaCl, 1 mM EDTA, 0.1% SDS.

1X washing buffer: 20mM Tris-HCl pH7.5, 0.1 M NaCl, 1mM EDTA, 0.1% Triton X-100.

1X ligation buffer: 20 mM Tris-HCl pH7.6, 25 mM Potassium acetate, 10 mM magnesium acetate, 10 mM DTT, 1 mM NAD, 0.1% Triton X-100.

6. Ligation in Streptavidin-coated PCR tubes

make a master solution and each tube contains 49 µl 1X ligation buffer
1 µl Taq DNA Ligase (40 U/ µl)

incubate at 60°C for 60 min.

wash each tube 1x with 1X washing buffer
1x with ddH2O

7. Elution of ligated products

add 50 µl ddH2O to each tube and incubated at 95°C for 5 min, chilled on ice, transfer the supernatant to a clean tube.

8. PCR set up

25 mM dNTPs	0.5 µl
10X buffer II (PEB)	2.5 µl

25mM MgCl2	1.5 µl
AmpliTaq Gold DNA Polymerase (5 Units/µl, PEB)	0.3 µl
Eluted (ligated) product (see above)	3 µl
Primer set (T3/T7/T7v, 10 µM each)	2 µl
ddH2O	l µl
Total volume	25 µl

PCR condition:
94°c 10 min
35 cycles of 94°C 30 sec
60°C 30 sec

and then
72°C 30 sec

CLAIMS

WHAT IS CLAIMED IS:

1. A method of detecting at least first and second target molecules in a sample comprising:

a. contacting said first and second target molecule with a composition comprising:

i. an amplification enzyme; and

ii. first and second target probes, said first and second target probes comprising:

a) a first and second bioactive agent, respectively, wherein said first and second bioactive agents specifically bind to said first and second target molecules, respectively;

b) a first and second adapter sequence, respectively, wherein said first adapter sequence identifies said first target molecule and said second adapter sequence identifies said second target molecule; and

c) at least first and second upstream universal priming sequences;

b. amplifying said first and second adapter sequences using said first and second universal priming sequences, wherein no ligation is performed, to form first and second amplicons, respectively;

c. detecting said first and second amplicons, respectively, to indicate the presence or absence of said first and second target molecules in said sample.

2. The method according to claim 1, wherein said first and second target molecules are selected from the group consisting of proteins and nucleic acids.

3. The method according to claim 1 or 2, wherein said first and second target molecules are nucleic acids.

4. The method according to claim 1, 2 or 3, wherein said first and second bioactive agents are first and second target specific domains that are substantially complementary to at least one domain of said first and second target nucleic acids, respectively.

5. The method according to claim 1, 2, 3 or 4, wherein said first and second target nucleic acids further comprise a first and second detection position, respectively, and said first and second target specific domains comprise a nucleotide at a readout position that is perfectly complementary to the nucleotide at said first and second detection positions.

6. The method according to claim 1, 2, 3 or 4, wherein said first and second upstream universal priming sequences are RNA Polymerase primers and said enzyme is an RNA Polymerase.

7. The method according to claim 6, wherein said RNA Polymerase primers are T7 primers.

8. The method according to claim 1, 2, 3 or 4, further comprising contacting said first and second upstream universal priming sequences with first and second chimeric RNA/DNA primers, wherein said amplification is by SPIA.

9. The method according to claim 1, 2, 3, or 4, wherein said first and second target probes further comprise first and second downstream universal priming sequences, wherein said first and second upstream universal priming sequences and said first and second downstream universal priming sequences flank said first and second adapter sequences, respectively.

10. The method according to claim 9 further comprising contacting said first and second upstream universal priming sequences and said first and second downstream universal priming sequences with first and second universal primers, wherein said amplification is by PCR.

11. The method according to claim 1 or 2, wherein said first and second target molecules are proteins.

12. The method according to claim 1, 2 or 11, wherein at least said first bioactive agent is an antibody.

13. The method according to claim 1, 2 or 11, wherein at least said first bioactive agent is a ligand.

14. The method according to claim 1, 2 or 11, wherein at least said first bioactive agent is an aptamer.

15. The method according to claim 1, 2 or 11, wherein said first and second upstream universal priming sequences are RNA Polymerase primers and said enzyme is an RNA Polymerase.

16. The method according to claim 15, wherein said first and second upstream universal priming sequences are T7 RNA Polymerase primers.

17. The method according to claim 1, 2 or 11, further comprising contacting said first and second upstream universal priming sequences with first and second chimeric RNA/DNA primers, respectively, wherein said amplification is by SPIA.

18. The method according to claim 1, 2 or 11, wherein said first and second target probes further comprise first and second downstream universal priming sequences, wherein said first and second upstream universal priming sequences and said first and second downstream universal priming sequences flank said first and second adapter sequences, respectively.

19. The method according to claim 1, wherein said composition further comprises nucleotides.

20. The method according to claim 1 or 19, wherein at least said nucleotides are labeled nucleotides.

21. The method according to claim 1, wherein said detecting comprises:

a. contacting said first and second amplicons with at least one substrate comprising first and second capture probes, wherein said first capture probes are complementary to said first adapters and said second capture probes are complementary to said second adapters; and

b. detecting said first and second amplicons on said at least one substrate.

22. The method according to claim 21, wherein said substrate comprises an array, said array comprising at least first and second capture probes immobilized on said substrate.

23. The method according to claim 22, wherein said array is an ordered array.

24. The method according to claim 21, wherein said substrate comprises at least a first and a second population of microspheres, wherein said first capture probes are immobilized on said first population of microspheres and said second capture probes are immobilized on said second population of microspheres.

25. The method according to claim 24, wherein said first and second amplicons are detected in a liquid array.

26. The method according to claim 25, wherein said first and second amplicons are detected by FACS.

27. The method according to claim 24, wherein said microspheres are randomly distributed on a second substrate comprising discrete sites.

28. The method according to claim 24, wherein said microspheres are applied to a mass spectrometer and the mass of said adapter sequence is determined to identify the presence of

a. contacting said plurality of target molecules with a composition comprising:
   i. an amplification enzyme; and
   ii. a plurality of target probes, each comprising:
      a) a bioactive agent, wherein each bioactive agent binds to a unique target molecule;
      b) an adapter sequence that identifies said unique target molecule that binds the bioactive agent; and
      c) at least one upstream universal priming sequence;
b. amplifying said adapter sequences using said at least one universal priming sequence, wherein no ligation is performed, to form a plurality of amplicons;
c. detecting said plurality of amplicons, to indicate the presence or absence of said target molecules in said sample.

30. The method according to claim 29, wherein said target molecules are selected from the group consisting of proteins and nucleic acids.

31. The method according to claim 29 or 30, wherein said target molecules are proteins.

32. The method according to claim 29 or 30, wherein said target molecules are nucleic acids.

33. The method according to claim 29 or 30, wherein said plurality of target molecules comprises at least 500 target molecules.

34. The method according to claim 29 or 30, wherein said plurality of target molecules comprises at least 1000 target molecules.

A Flow Chart for Array-based Detection of Gene Expression

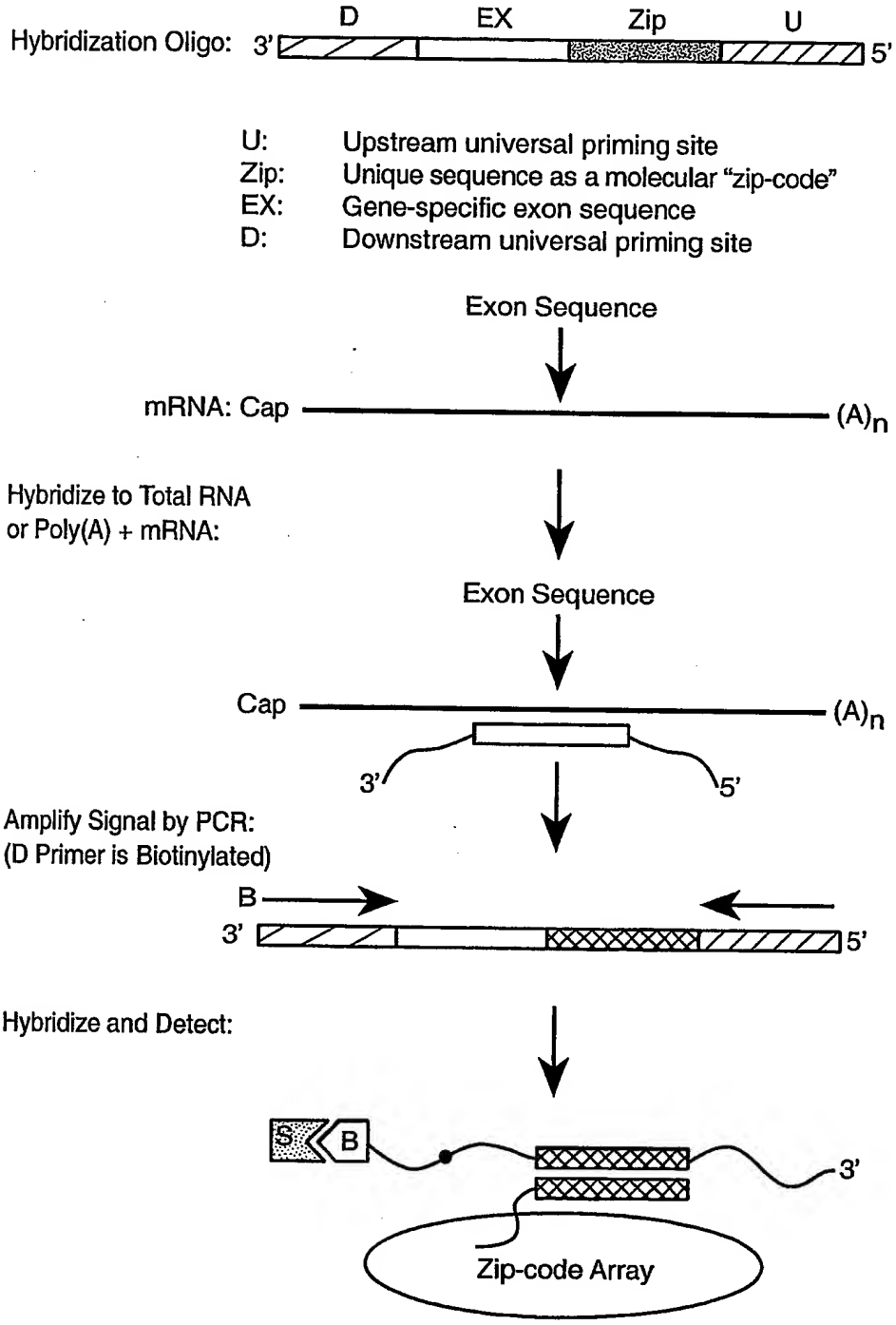


FIG._1

A Flow Chart for Array-based Detection of RNA Alternative Splicing

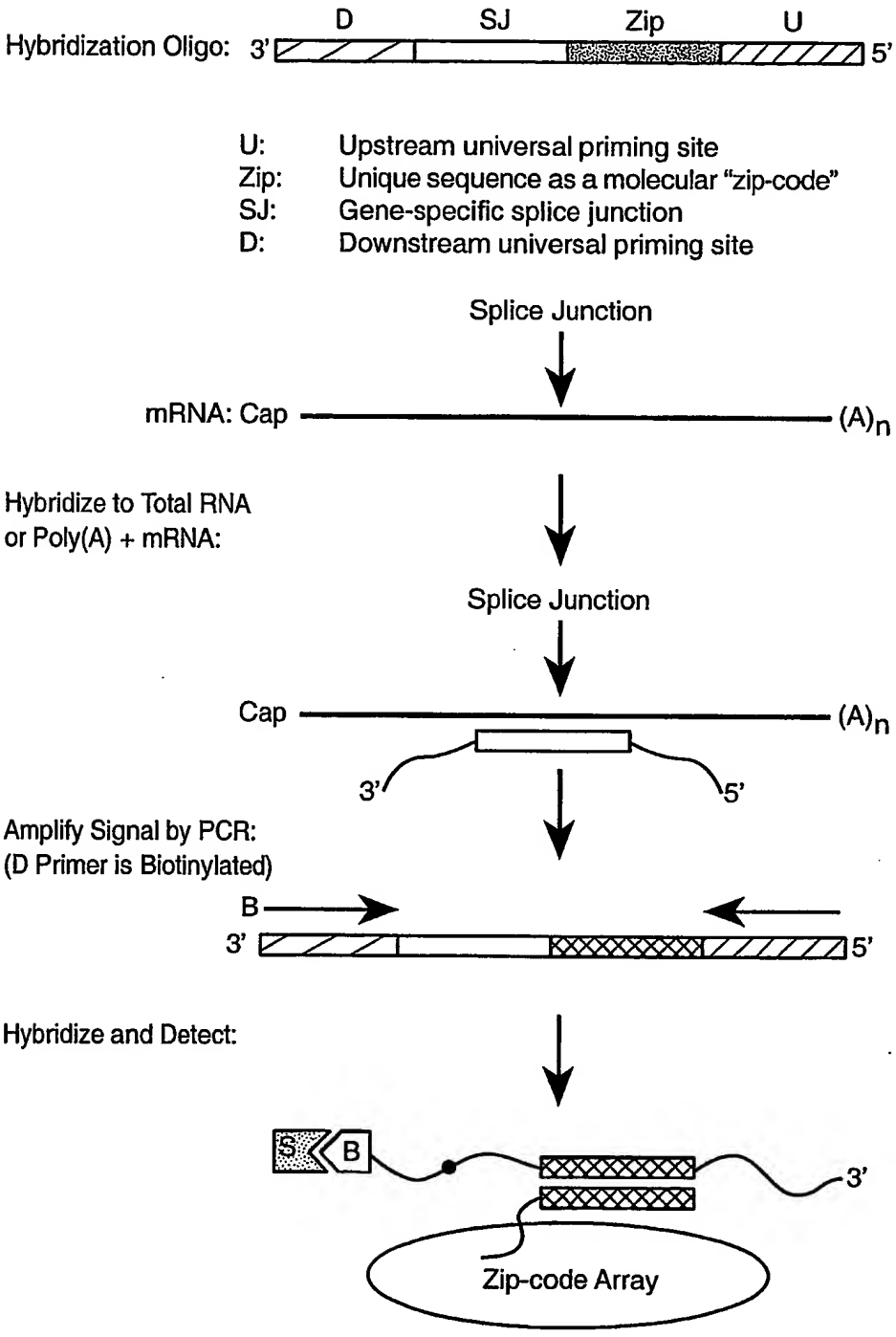


FIG._2

Genome-wide Gene Expression Profiling Using Oligo-ligation Strategy

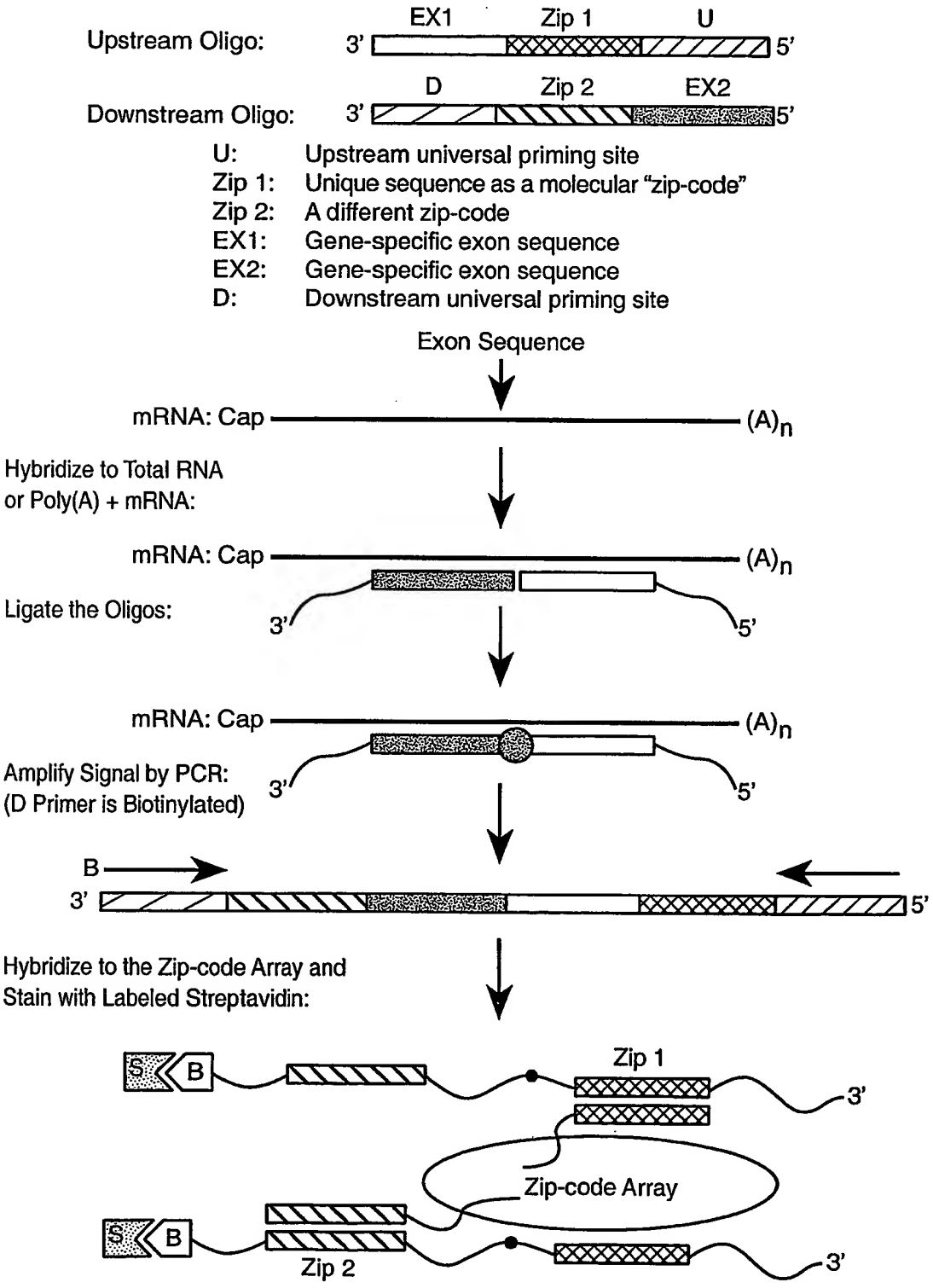


FIG._3

Genome-wide RNA Alternative Splicing Monitoring Using Oligo-Ligation Strategy

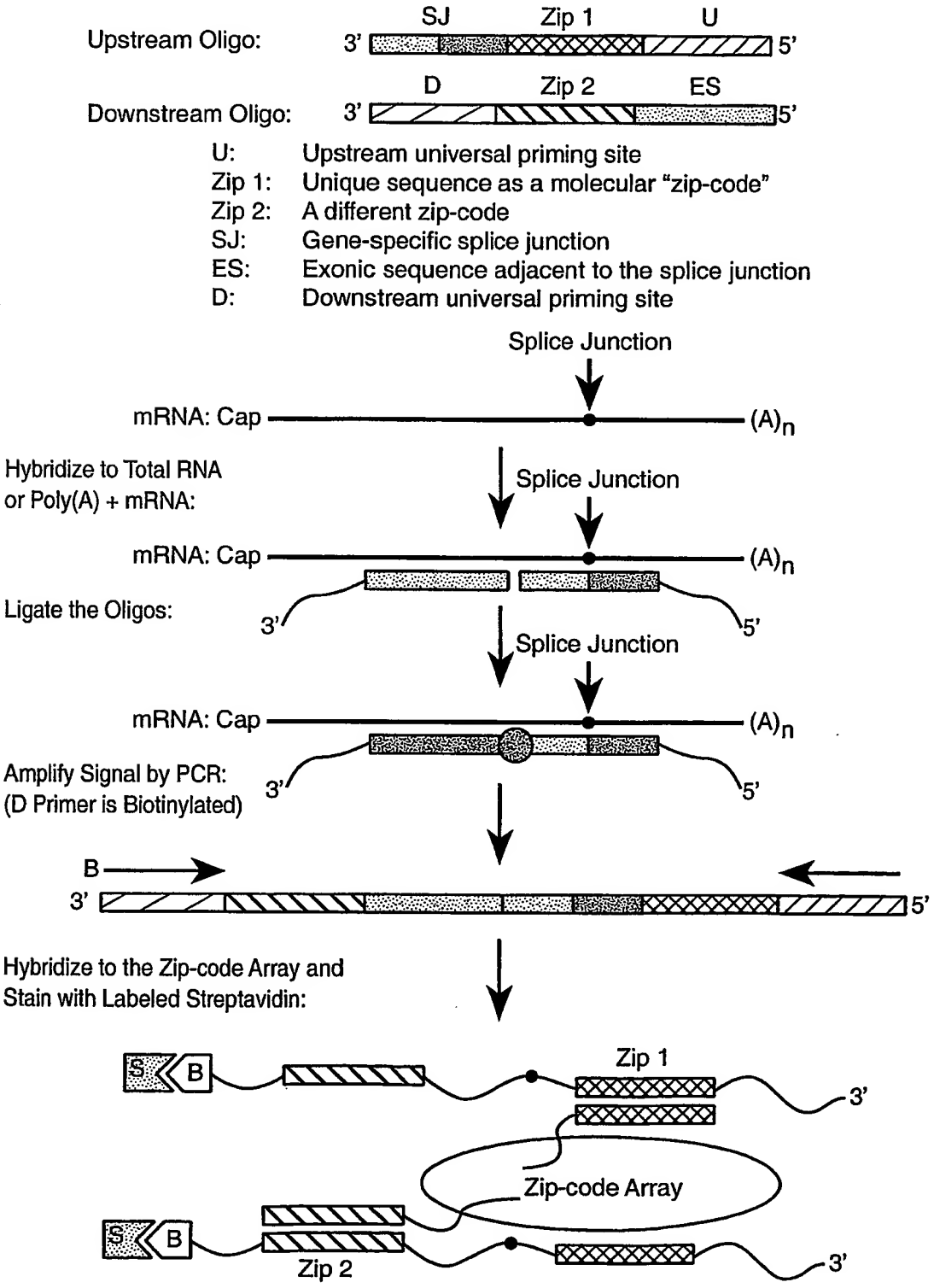


FIG._4

Direct Genotyping Using a Whole-genome Oligo-ligation Strategy

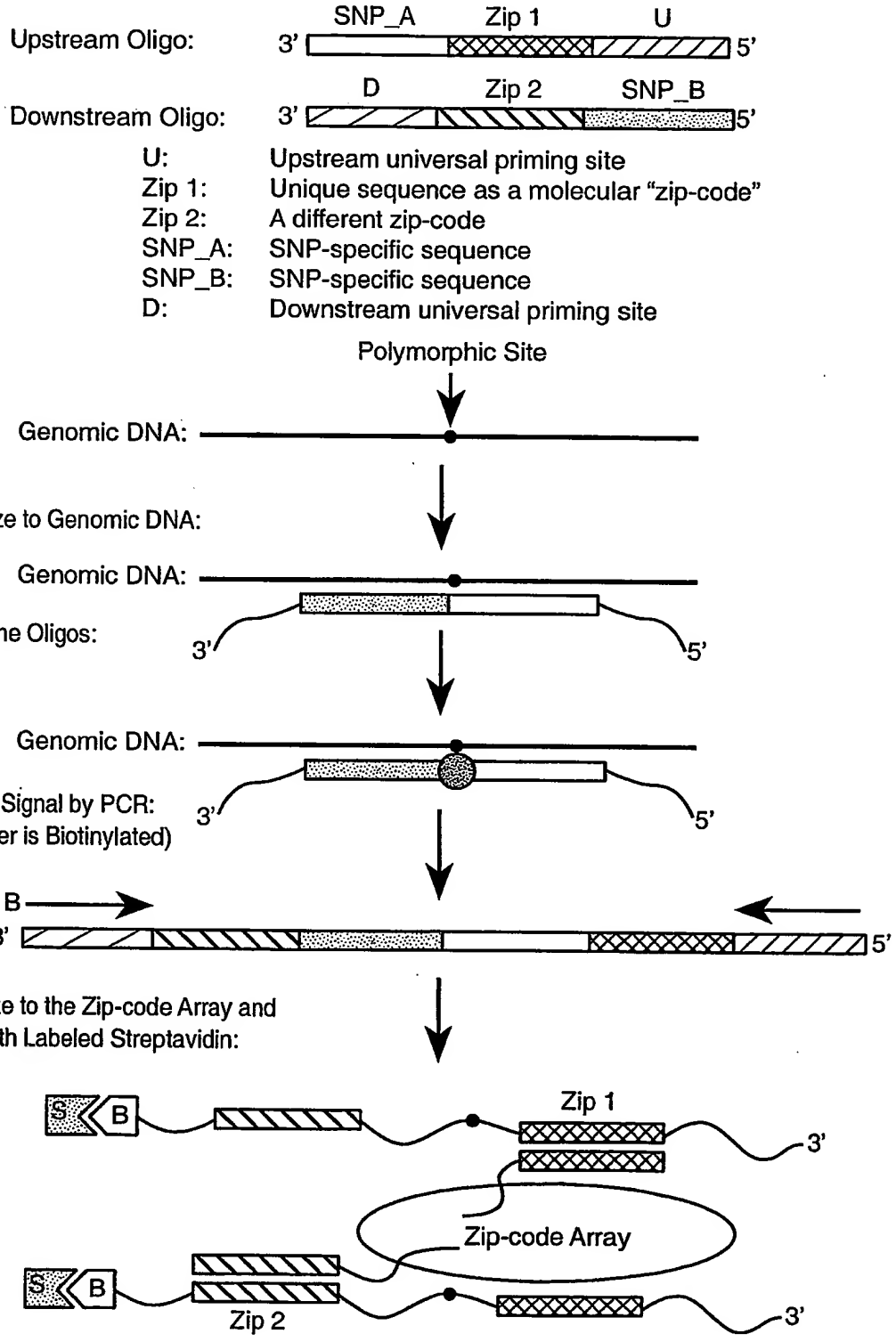


FIG._5

Whole-genome Oligo-ligation Strategy

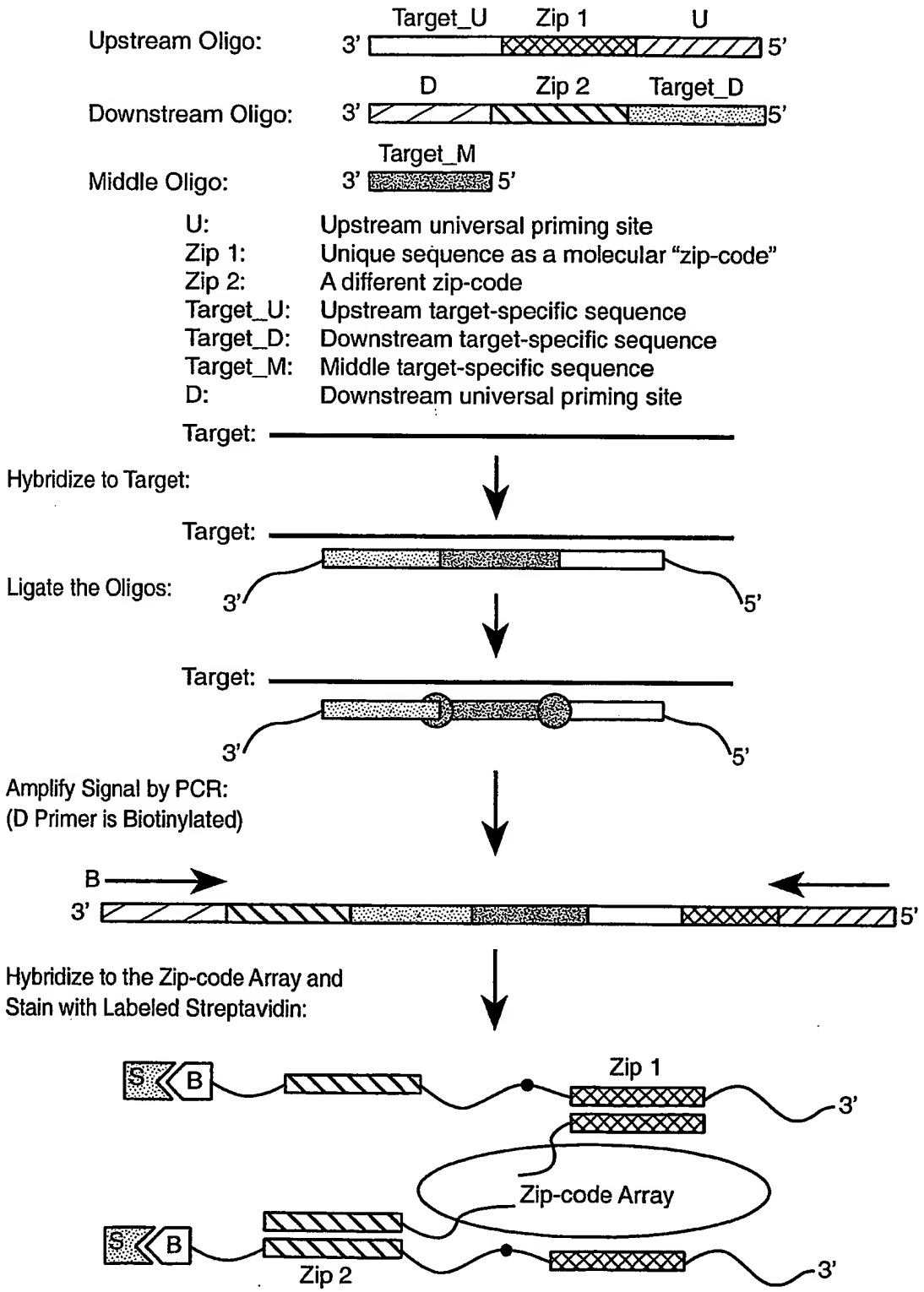


FIG._6

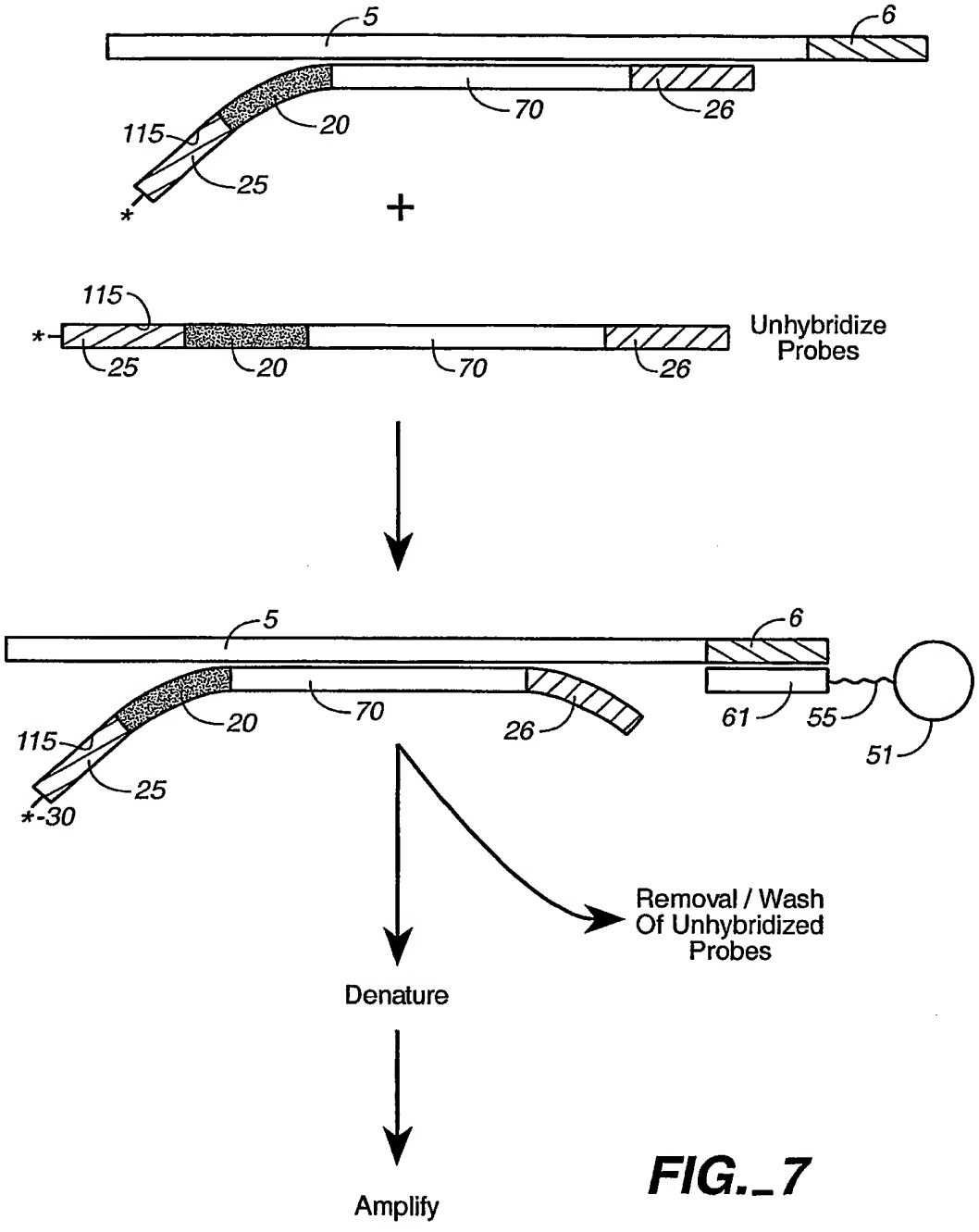

5
6
115
20
70
26
25
*
+
115
*
25
20
70
26
Unhybridize Probes
5
6
20
70
26
61
55
115
25
*-30
51
Removal / Wash Of Unhybridized Probes
Denature
Amplify

FIG._7

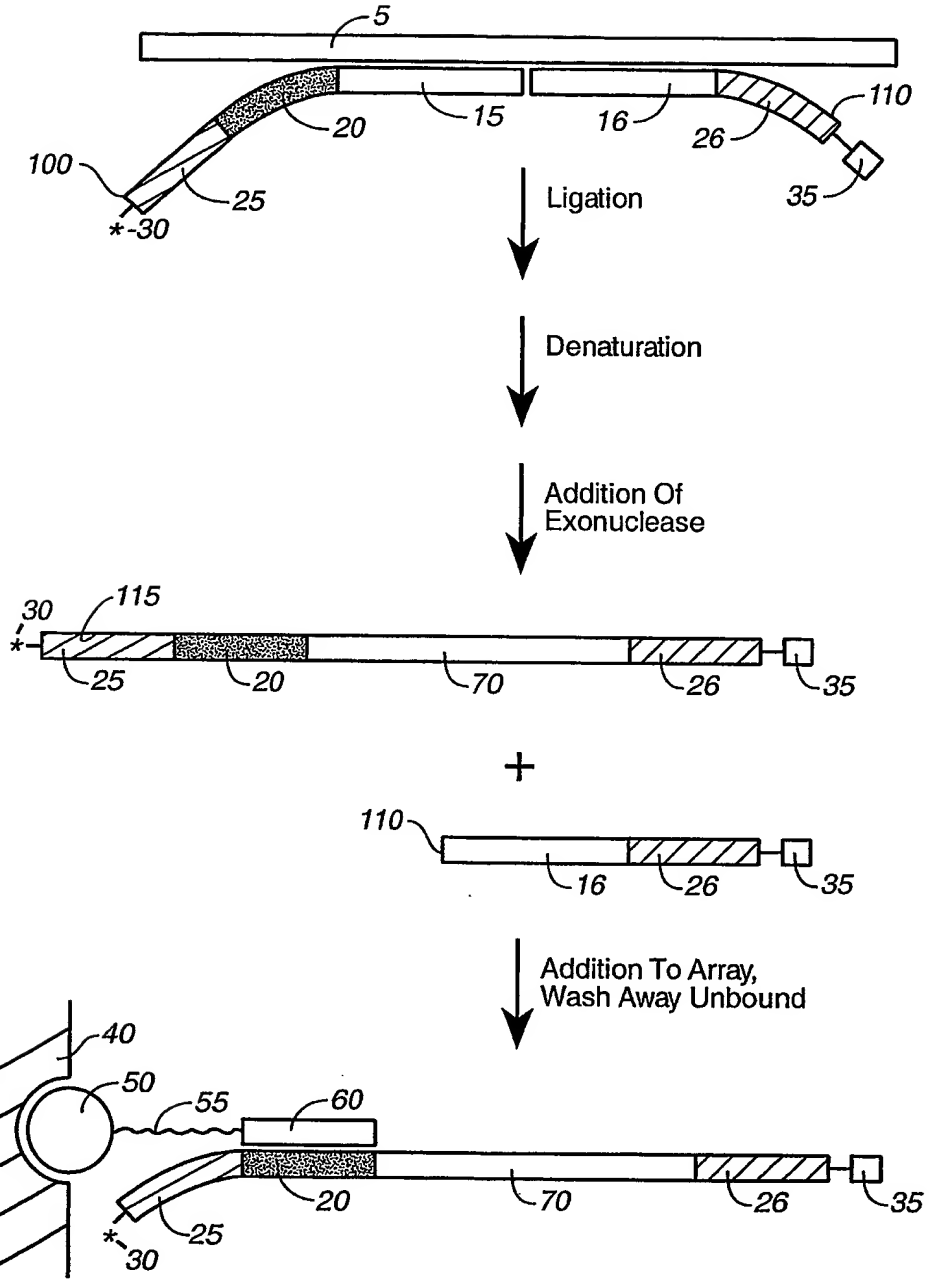

5
110
20
15
16
26
100
25
35
*-30
Ligation
Denaturation
Addition Of
Exonuclease
30
115
25
20
70
26
35
+
110
16
26
35
Addition To Array,
Wash Away Unbound
40
50
55
60
20
70
26
35
25
*-30

FIG._8

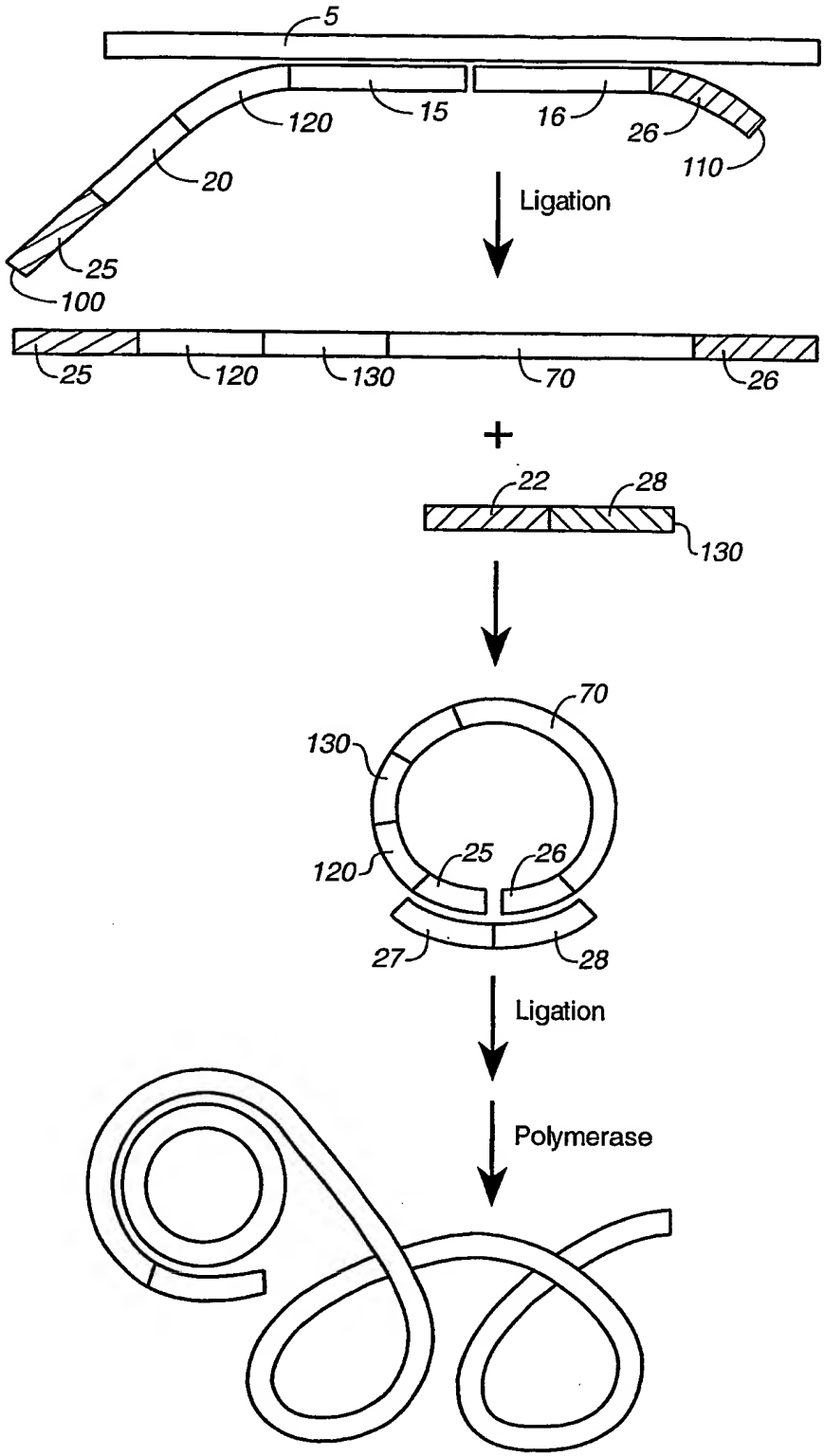

5
120
15
16
26
110
20
Ligation
25
100
25
120
130
70
26
+
22
28
130
70
130
120
25
26
27
28
Ligation
Polymerase

FIG._9

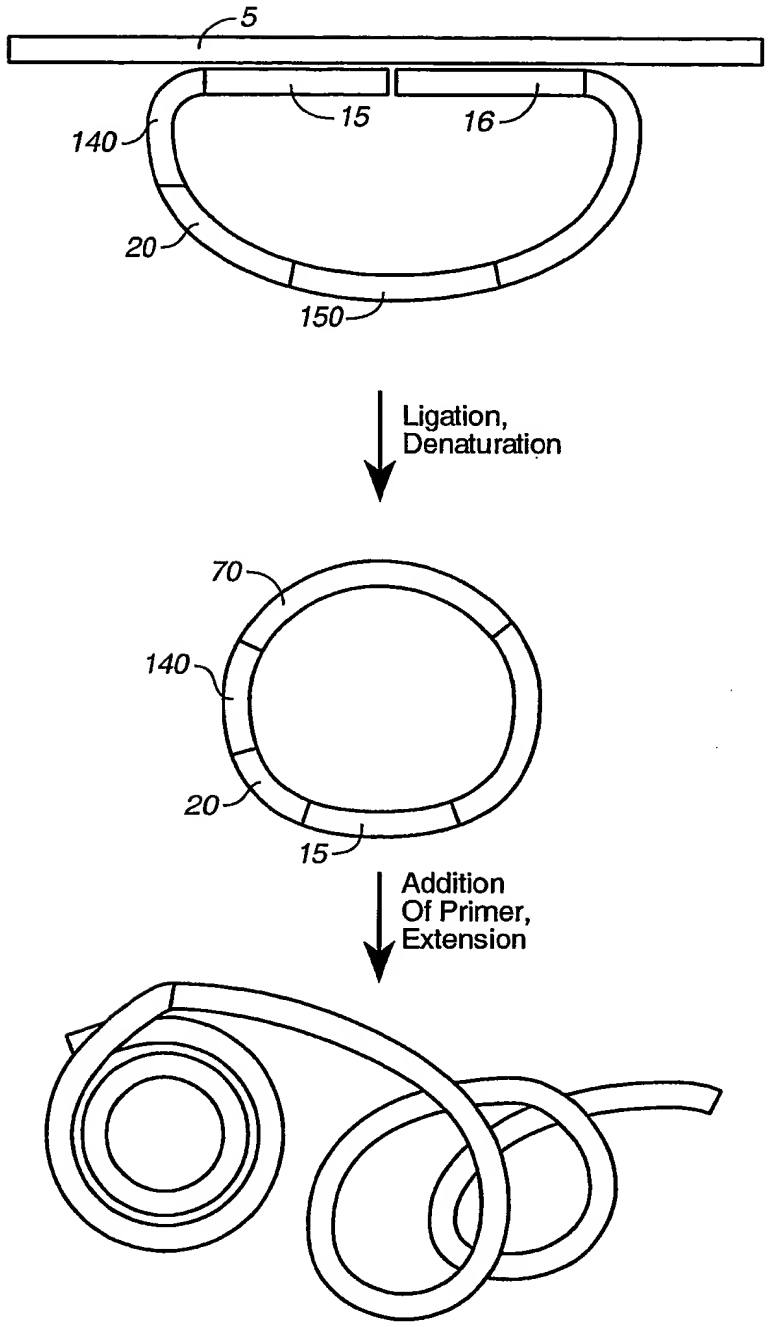

5
15
16
140
20
150
Ligation,
Denaturation
70
140
20
15
Addition
Of Primer,
Extension

FIG._10

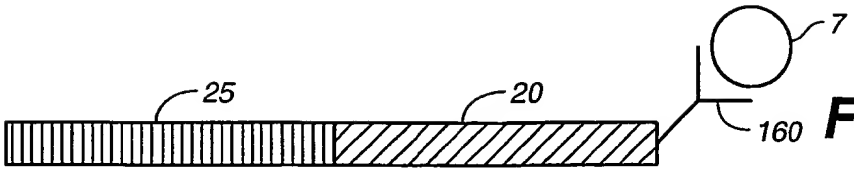


FIG._11A

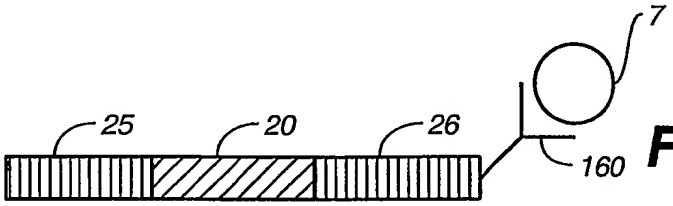


FIG._11B

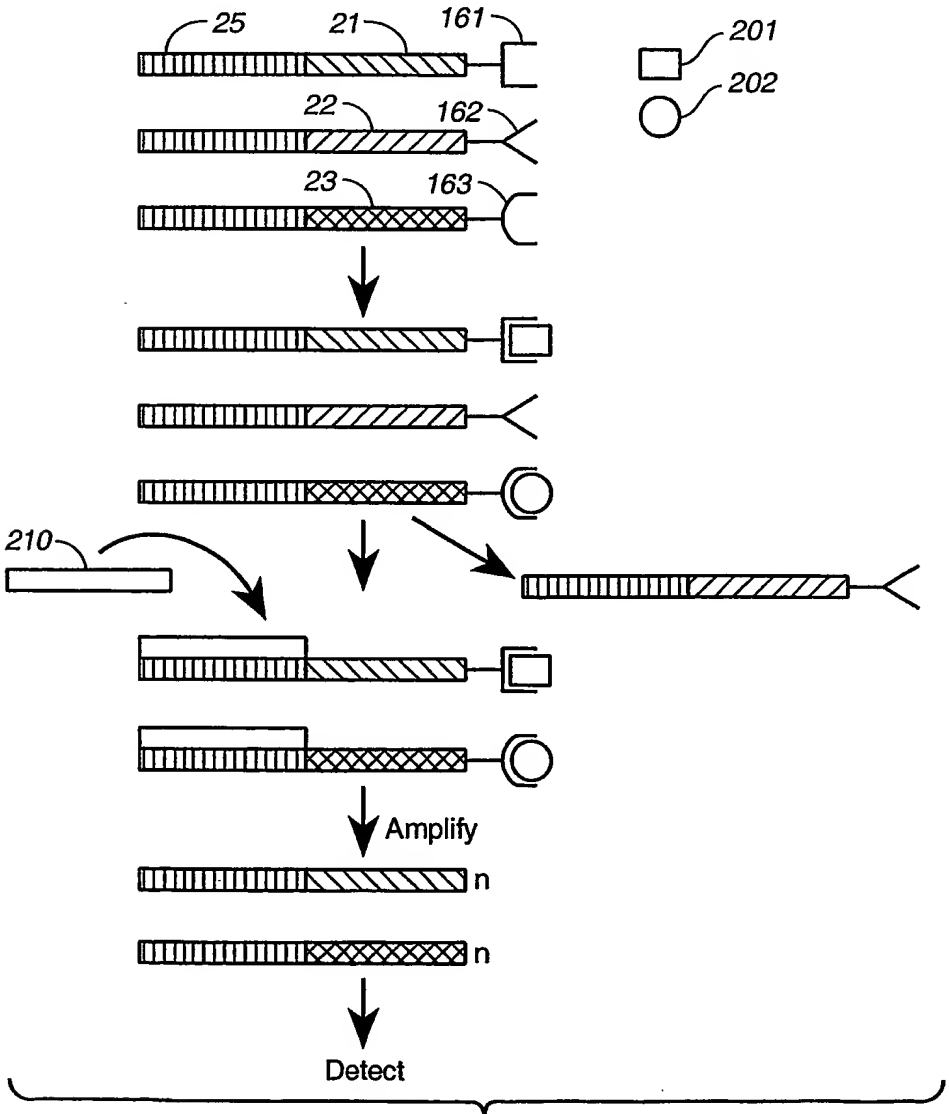


FIG._12

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.